

Real-time Capability of AR-based Control for Overhead Cranes in Intralogistics

Anna-Lena Fritze¹

Eike Lorenz²

Carsten Schmidt¹

Benjamin Küster²

Ludger Overmeyer^{1,2}

¹Institute of Transport and Automation Technology (ITA), Leibniz University Hannover ²Institut für Integrierte Produktion Hannover gGmbH (IPH)

Automation of intralogistics processes offers significant potential for improving industrial efficiency. To date, there are few established approaches for transporting bulky loads using overhead cranes. A human-machine interface incorporating augmented reality is intended to make operation more intuitive and safer in this approach. To ensure the applicability of the proposed operator assistance system in complex environments, a high degree of responsiveness is required. The demonstrator utilizes ROS 2 as the communication framework between the computing units, enabling the direct temporal association of operator commands with the measured crane accelerations. The results presented in this paper demonstrate that the latency for processing and executing gesture-based operator inputs is below 60 ms and can therefore be considered capable of real-time processing for human perception. Watchdog timers are implemented to ensure system safety in the event of network disruptions according to DIN EN IEC 61508. Conducting various test series the robust intervention of the timers is shown, thereby validating the stability of the novel interface making it suitable for further automation of the system.

[Keywords: Augmented Reality, Overhead Crane Assistance, Real-time Capability, Distributed Systems]

Die Automatisierung intralogistischer Prozesse bietet erhebliches Potenzial zur Steigerung der industriellen Effizienz. Bislang gibt es nur wenige etablierte Ansätze für den innerbetrieblichen Transport sperriger Lasten mit Brückenkranen. Eine Mensch-Maschine-Schnittstelle, die Augmented Reality nutzt, soll bei diesem Ansatz die Bedienung intuitiver und sicherer gestalten. Um die Anwendbarkeit des vorgeschlagenen Bedienerassistenzsystems in komplexen Umgebungen zu gewährleisten, ist ein hohes Maß an Reaktionsfähigkeit des Systems erforderlich. Als Kommunikationsebene zwischen den Rechnern wird für den Demonstrator ROS 2 implementiert, das ermöglicht das direkte zeitliche Kopplern der Befehlseingaben an die Beschleunigungen des Kranes. Die in

diesem Beitrag vorgestellten Ergebnisse zeigen, dass die Latenzzeit für die Verarbeitung und Ausführung gestenbasierter Bedienereingaben unter 60 ms liegt und somit für die menschliche Wahrnehmung als echtzeitfähig angesehen werden kann. Zur Gewährleistung der Systemsicherheit bei Netzwerkstörungen entsprechend der DIN EN IEC 61508 werden Watchdog-Timer implementiert. In verschiedenen Testreihen wird das robuste Eingreifen der Timer nachgewiesen, wodurch die Stabilität der neuartigen Schnittstelle bestätigt und die Eignung für die weitere Automatisierung des Systems gezeigt wird.

[Keywords: Augmented Reality, Assistenzsystem für Brückenkranen, Echtzeitfähigkeit, Verteilte Systeme]

1 INTRODUCTION

The automation of intralogistics processes is an important area of research in respect to growing demands for faster throughput times and increased efficiency, as well as increasingly stringent safety regulations. In the context of Industry 5.0, the focus is on human as a value-adding factor. Automation solutions should be people-centred and collaborative [1]. This results significant challenges, particularly for automated crane systems. In context of intralogistics, overhead cranes are primarily used to transport bulky and heavy loads, which consequently entails a high safety risk [2]. The sources of risk for workplace accidents include lifting loads without proper control leading to pendel behaviour and misjudging situations due to limited visibility as well as insufficient safety distances. Due to the high safety requirements for automation, crane systems in human-occupied workspaces are operated exclusively by human operators. For internal transport, they are typically controlled from the floor via radio or a wired connection. Fully automated systems exist only in isolated areas with a deterministic initial state.

The research underlying the investigations presented in this paper examines the potential of Augmented Reality

(AR)-based control systems. The goal is to reduce operating errors and enhance safety through perspective-correct warning messages, similar to the investigations of [3]. The user interface proposed in this paper provides both visualization capabilities and direct interaction with the crane control system. By controlling the crane via natural user interfaces, the controllers cognitive load is reduced and the effectiveness of the transport processes increases [4]. Figure 1 illustrates a conceptual future interface that enables the intuitive selection of target locations within the operating environment.

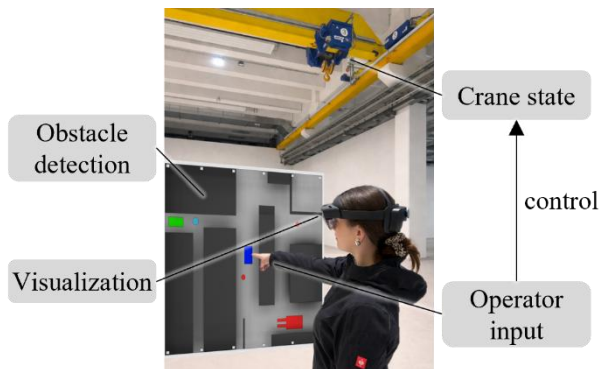


Figure 1. AR-based human-machine-interface enabling intuitive interaction

The availability of information about the dynamic environment enables the implementation of assisting functionalities including route planning and collision avoidance progressing to fully automated transport. Such human-machine interaction requires access to a variety of system state information. To acquire this information, a sensor infrastructure needs to be deployed throughout the workspace of the crane. The infrastructure comprises camera and LiDAR sensors mounted either statically within the environment or dynamically on the crane trolley, as well as the AR headset serving as the human-machine interface (HMI). To enable low-latency processing of resulting sensor data, sensor outputs are processed in a decentralized manner across distributed computing systems and subsequently fused on a central computing unit (CCU). The real-time transmission of this information is essential from both a safety perspective and with regard to system usability.

The primary research objective of this paper is therefore to determine the system latencies and system safety for real-time use in a real production environment focusing on the control signals. The control signal is digitized by an AR-platform developed in Unity, which allows the operator to interact with the crane control system using gestures. The signal is then processed further in a Robot Operating System (ROS), here a ROS 2 network, and finally transmitted as a radio signal to the crane control system via a microcontroller-driven relay circuit. To validate the proposed approach, a crane demonstrator is implemented at the Institute of Transport and Automation Technology, where a series of experiments is conducted.

2 RELATED WORK

The application of AR as an automation technology and as an interface between the operator and the crane has been subject of extensive research. In this context, a distinction must be made between assistance systems and AR-based control approaches. Latter receives considerably less attention in the existing literature, while the influence of latency on signal transmission is of considerably greater relevance.

HMI. The optimization of the interface between operator and crane has been subject of numerous research publications [5]. In 2012, the reduction of collisions through natural gesture-based trajectory specification has been successfully demonstrated. The proposed interface relies on conventional computer vision using a camera mounted on the crane trolley, which limits both the transparency of system decisions and the operator's interaction capabilities [6]. Extended Reality (XR) as an operator interface enables the use of various input devices and natural user interaction like gestures to improve intuitiveness [7]. Particularly in the context of remote handling, fully automated approaches are proposed in which the operator only specifies target locations while the crane autonomously executes the motion using rule-based control algorithms [8]. Furthermore, the interoperability of Unity as an AR development platform and ROS within an XR application context is demonstrated [9].

AR control. In the context of digital twins, the potential inherent in virtual representations is increasing significantly. Based on a digital twin, [10] investigates a mixed reality interface employing a HoloLens 1 for the control of an overhead crane. A positioning accuracy of 10 cm is achieved, however, temporal latencies within the system are not considered as part of the study [10]. Likewise, [11] investigates an AR-based interface for the remote operation of a scaled tower crane using a digital twin. In this case, a smartphone serves as user interface. The end-to-end latency of the system is not evaluated and the digital twin is updated at a rate of 2 Hz. Similar investigations conducted by [12], whose proposed system exhibits with a latency within the AR controller of approximately 250 ms.

XR latency. Research on latency in XR environments has largely focused on approaches based on virtual reality (VR). In their VR-based industrial control systems, [11, 12] specify frame rate requirements corresponding to frame times below 30 ms. For control inputs, latencies of up to 200 ms are considered comparable to those of conventional remote-control interfaces. Furthermore, a maximum overall latency of $T_{total} \leq 110 \text{ ms}$ has been identified as a requirement for the real-time operation of machinery through an XR interface [15].

The AR assistance systems reviewed in the literature primarily exhibit latencies of approximately 100 ms for visual feedback and 200 to 500 ms for cloud- or middle-

ware-based communication [9]. For basic digital twin applications, update intervals in the range of 0.5–1 s have been reported [11]. The delays caused by the high demands placed on data processing result in a loss of real-time capability. For AR glasses, these latencies are not user-friendly and lead to an immediate reduction in safety. However, none of the investigated AR-based crane studies explicitly quantifies the complete end-to-end latency from an operator command issued through the AR interface to the resulting physical crane motion.

3 METHODOLOGY

To implement the AR-based control system, the interfaces between overhead crane and AR glasses must be designed to support human-machine communication. Focussing on continuous, low-latency data exchange.

3.1 NETWORK STRUCTURE

The presented concept imposes various requirements on the system network. Figure 2 illustrates the designed network architecture, with the end-to-end control path traversed by operator-issued control signals highlighted in red. ROS 2 is implemented as middleware for communication between the devices. The interface to the crane operator, who carries the AR device, is wireless to ensure freedom of movement. As HMI hardware Microsoft's AR glasses HoloLens 2 are used and connected directly to the ROS network via a Unity remote transmission with an integrated ROS-TCP connection. This enables direct publishing and subscribing of topics from the ROS network. A CCU serves as the network host and time master for synchronization of the system time, as well as an interface for additional sensors used in the research project. Serial interfaces are used between sensors, microcontroller units (MCU), and the CCU to ensure secure data transmission. This paper examines a setup, where the control commands to the overhead crane are implemented using a modified, relay-driven radio remote control. The control states are then set via serial communication with an Arduino MEGA 2560 based on the inputs interpreted in Unity.

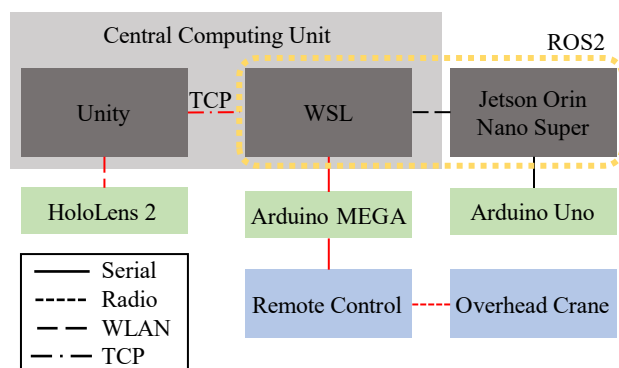


Figure 2. Network Architecture including ROS 2 (Red: End-to-End Control Path)

To track the crane's movement, a Jetson Orin Nano Super Developer Kit combined with an Arduino Uno Q is mounted on the crane trolley. The devices collect and process data from mobile sensors and transmit the results to the CCU via the ROS network, where further processing takes place.

3.2 INVESTIGATION OF REAL-TIME CAPABILITY

Particularly with regard to the safety requirements of automated and distributed systems, latencies are of great importance. As the control chain has been extended by several intermediate interfaces, an increase in response time of the control system is expected. The end-to-end control path shown in Figure 2 illustrates the various data transmission technologies involved, each representing a potential source of failure. Figure 3 highlights the significance of these potential failure sources within the process through which a control command initiated by the crane operator is transmitted.

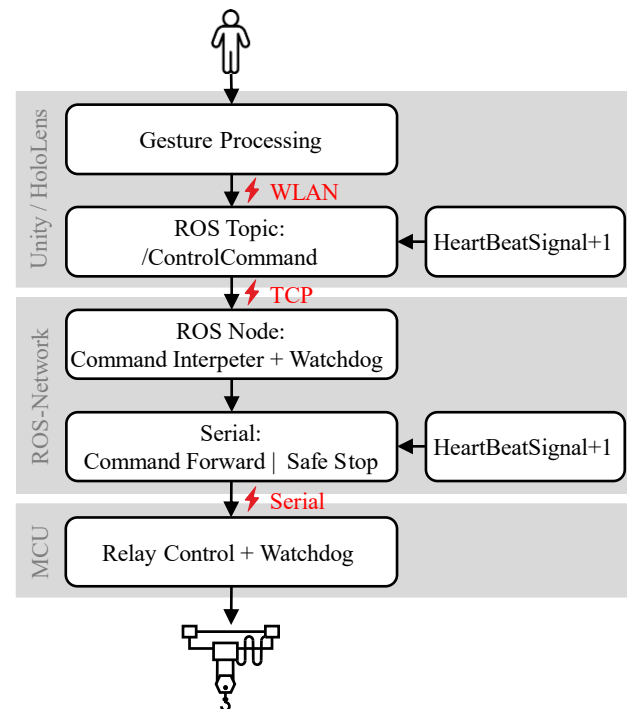


Figure 3. Propagation of the control command through the system components and potential failure points within the communication chain (red)

The HoloLens serves as the primary HMI converting the operator's natural gestures into digital control signals. To maximize computational performance, a remote connection is established between the HoloLens and the Unity program on the CCU. In the event of a wireless network failure, this connection is interrupted, preventing the operator from issuing further control commands. Such a condition is considered safety-critical and must result in an immediate stop of crane operation. Under normal network conditions, the user input processed by Unity is published

as a topic within the ROS network via the TCP bridge. Similar to a wireless network failure, a disruption of this connection would prevent the CCU from receiving operator commands. To ensure signal integrity, watchdog timers are deployed on the relevant system instances to continuously monitor the data flow within the network. The monitoring process operates at a frequency of 100 Hz. If the connection to Unity is interrupted, heartbeat signals are no longer received, causing the system to transition into a safety stop state after a predefined watchdog timeout. During regular operation, the control commands are transmitted via the ROS node to the MCU through a serial communication link. The MCU then updates the states of the radio remote control by actuating the corresponding relays. To ensure the reliability of the serial communication, an additional watchdog mechanism is implemented. In the event of a serial communication failure, it likewise forces the system into the safety stop state.

The aim of this paper is to validate the system's responsiveness within the framework of a defined real-time capability. In addition, system safety is evaluated in accordance with DIN EN IEC 61508 to ensure that the control system is usable for future investigations [16]. In the following, the response time describes the time required to react to an external event. This includes both the ability to execute control commands planned by the operator and the ability to detect disruptions in the network architecture and carry out appropriate system responses. In this case, the definition of real-time capability refers to the reliable execution of calculations and commands within a specified time frame. For the watchdog timers, a response time of 250 ms is set, which approximates human visual reaction time [17], aiming to simulate a human reaction to an error. The overhead crane used for this demonstrator purpose provides two controllable speed stages for each direction of motion. According to the data sheet, the maximum travel speed for bridge travel motion is $v_{\max} = 25$ m/min. For the defined watchdog time, this corresponds to an active travel distance of 10.416 cm when the communication link is interrupted. This value is considered acceptable, provided that the operator maintains the legally required safety distances of 0.5 m [18]. The evaluation criterion is the time at which the relays are set to the required state, based on the commands actively generated within the system. Hereby, the delays in the crane control system between the transmission of the radio signal and the crane's actual movement are not part of the investigation.

4 EXPERIMENTAL SETUP

To conduct stress tests and analyse latencies in the transmission pipeline, a demonstrator is built. The setup allows acceleration tracking of the cranes' motion axis and sending control commands via an augmented operating panel. A series of measurements is carried out to evaluate the system's responsiveness and robustness.

4.1 HARDWARE INTEGRATION

For experimental evaluation, the proposed network architecture was integrated within an existing crane system. The individual system components are shown in Figure 4.

1. HoloLens 2: The HMI communicates with the CCU via a 5 GHz wireless network.
2. The remote-control transmitter: Replacing the push buttons of a commercial remote control by relays, enables actuating the crane by an Arduino MEGA.
3. Sensor processing unit: Integrated into ROS network through the Jetson platform, it performs IMU sensor data processing locally at the crane trolley.

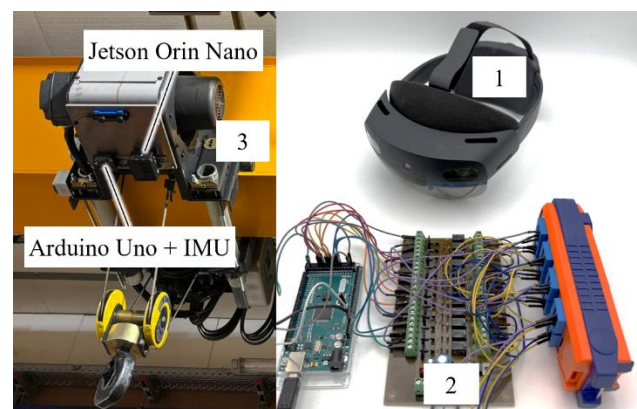


Figure 4. Demonstrator setup

For the investigations of the crane motion a 6-axis IMU sensor (LSM6DSOXTR) is implemented to monitor acceleration behaviour. The values are analysed with a measurement range of ± 2 g to maximize measurement resolution while sampling acceleration data at 833 Hz. An offset correction was performed using 100 measurement samples to eliminate the effects of gravitational acceleration and sensor tilt on the recorded acceleration data.

4.2 HUMAN-MACHINE-INTERFACE

By using AR glasses as an HMI, user input can be provided in various ways. The use of the OpenXR plugin and the MixedRealityToolkit enables the positioning of interaction elements relative to the operator. The graphical user interface developed for these studies is modelled according to a conventional remote control, making it intuitive even for experienced crane operators (Figure 5). The primary advantage of the virtual panel is that the user interface can be adapted to the specific situation [19]. For example, the interface used in these experiments includes labels for the buttons that correspond to their orientation within the factory hall. For inexperienced crane operators, both the remote-control layout and spatial orientation within the operating environment may present significant challenges, which can lead to an increased risk of operator errors.

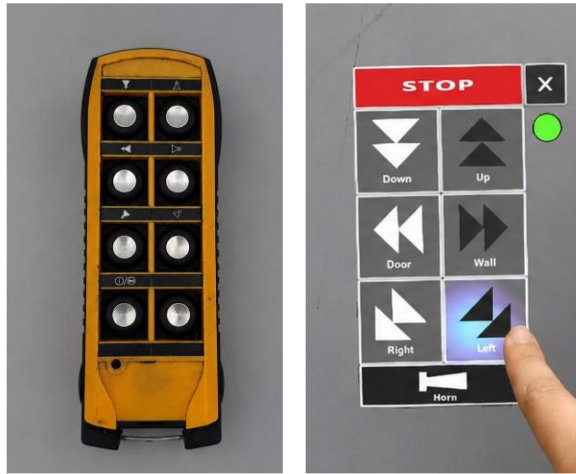


Figure 5. Human-Machine-Interface (Left: Commercial radio remote control; Right: AR-based holographic interface)

In addition, the operator can hide the implemented virtual remote control at any time. This allows him to use both hands, without having to set the control unit aside during phases when the crane does not need to be controlled. Large joystick remote controls, in particular, can quickly get in the way or restrict the operator's freedom of movement.

A virtual status indicator is provided to display the current connection state, thereby informing the operator about potential system disturbances. However, the emergency stop function, which has not yet been implemented, must be considered a critical limitation. Under normal operating conditions, the stop button can be used to halt crane motion. Although this functionality is executed in a dedicated thread, unforeseen deadlocks and other network-related failures cannot be excluded. In accordance with DIN EN ISO 13850, an emergency stop must enable the person actuating the switch to bring all crane movements to an immediate stop [20]. While the implemented safety

mechanisms reduce the risks associated with system failures, they cannot guarantee the level of reliability required for safety-critical operator intervention. Future work will therefore investigate the integration of an independent mechanical emergency stop device, for example directly attached to the AR headset.

4.3 EVALUATION AND RESULTS

To evaluate latencies, the acceleration profiles recorded by the LSM6DSOXTR are analysed and the ROS event processing tracked across the system instances. The measurement data of the acceleration signal is pre-processed to analyse the crane's acceleration. First a median filter with a window size of three samples is applied to suppress outliers. Subsequently, a fourth-order Butterworth low-pass filter with a cut-off frequency of 5 Hz is applied using zero-phase forward-backward filtering to reduce high-frequency noise without introducing phase distortion.

4.3.1 REAL-TIME-CAPABILITY

As defined in Section 3.2, the user input processed by the HoloLens passes through several network instances before being converted into a control command for the crane motors. According to (1), the end-to-end processing speed depends on latencies between the individual instances.

$$T_{\text{total}} = T_{\text{Unity} \rightarrow \text{ROS}} + T_{\text{ROS} \rightarrow \text{MCU}} + T_{\text{MCU} \rightarrow \text{Motor}} \quad (1)$$

Figure 6 shows the average times at which the control signals are processed in Unity, ROS and by the Arduino, as well as the resulting acceleration behaviour of the crane. To improve clarity, the release of the control button is hereafter referred to as an active stop command. From a digital signal processing perspective, both Unity inputs are equivalent. In all figures, the time axes are referenced relative to the triggering control event in order to emphasize the cor-

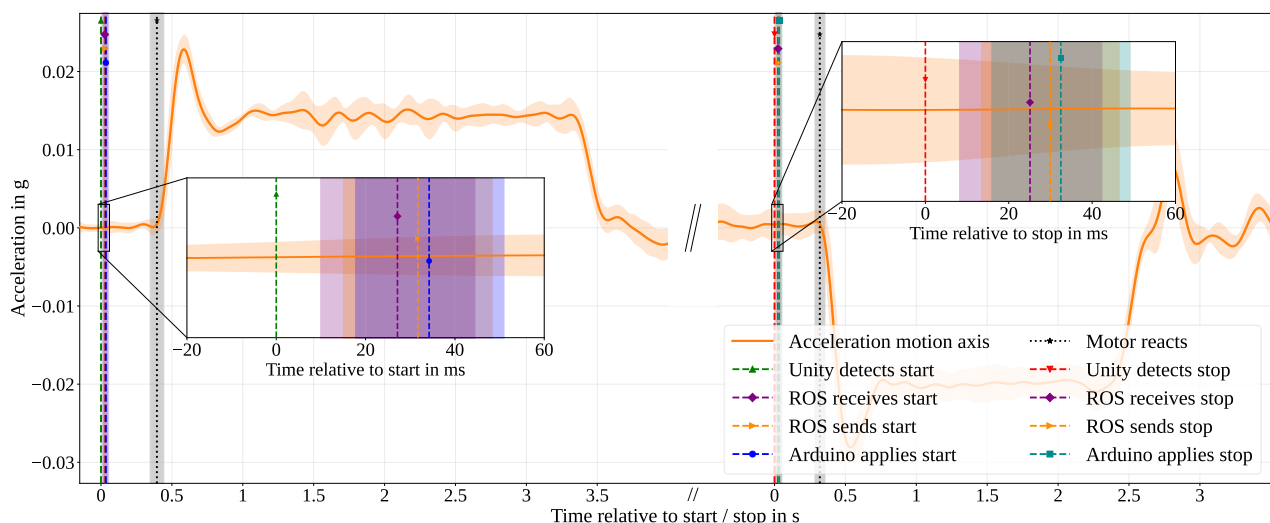


Figure 6. Signal processing of an AR-generated control command (Left: Start signal; Right: Stop signal) in the network instances and the corresponding acceleration behaviour of the crane

responding system responses. The study examines the system's response to a start signal from a stationary state, as well as a subsequent stop signal following movement at a constant speed. The experiment is repeated 15 times, thus, the average processing times for the control signals are $T_{\text{total,start}} = 392.260 \pm 45.701$ ms and $T_{\text{total,stop}} = 319.508 \pm 35.579$ ms. For data processing within the presented network, the results for start and stop command can be summarized with the actuator response considered separately.

The average signal processing time between HoloLens/Unity and the ROS network is:

- $T_{\text{Unity} \rightarrow \text{ROS}} = 26.120 \pm 16.833$ ms

with a minimum duration of 1.008 ms and a maximum of 58.632 ms. Followed by the ROS-side processing and serial transmission of the control command to the MCU takes place in:

- $T_{\text{ROS} \rightarrow \text{Unity}} = 7.244 \pm 2.686$ ms

The minimum $T_{\text{ROS} \rightarrow \text{MCU}}$ -latency is 3.504 ms, and the maximum 12.283 ms. The longest delay from the input of a control command by the operator to the actual change in the crane system is the time between the relays being activated and the actual crane movement:

- $T_{\text{MCU} \rightarrow \text{Motor,start}} = 360.034 \pm 45.458$ ms
- $T_{\text{MCU} \rightarrow \text{Motor,stop}} = 287.007 \pm 25.119$ ms

In consideration of the total delay in the command processing process, this means that the latency introduced by the new control method is comparatively low and well below the range perceptible to the user.

In particular, $T_{\text{ROS} \rightarrow \text{MCU}}$ within an acceptable range and behaves deterministically, as evidenced by its low standard deviation. Larger fluctuations can be identified for the HMI. The HoloLens typically operates at a frame rate of 60 fps. During testing, occasional frame gaps ranging from 20 ms to 80 ms are observed, which also explains the fluctuating delays. Considering the case with the greatest latency that occurred for the safety-critical stop command, we have a maximum unintended transfer of the load from (2):

$$\begin{aligned} \Delta x_{\text{max}} &= v_{\text{max}} \cdot (T_{\text{Unity} \rightarrow \text{ROS,max}} + T_{\text{ROS} \rightarrow \text{MCU,max}}) \quad (2) \\ &= 25 \frac{\text{m}}{\text{min}} \cdot 70.915 \text{ ms} \approx 0.0295 \text{ m} \end{aligned}$$

Taking into account pendulum motions and the achievable positioning accuracy of two-stage overhead cranes also compared with [10], the results can be considered suitable for real-time applications and not safety critical.

4.3.2 SAFETY

To assess system safety, controlled communication outages are introduced and the responsiveness of implemented safety mechanisms evaluated. As discussed in Section 3.2, a maximum response time of 250 ms is specified as the real-time requirement for the watchdog timers. System behaviour is analysed for 15 measurement series for each interruption. All events within the communication chain are referenced relative to the time of the disturbance.

Wlan interrupt. An interruption of the wireless network connection to the HoloLens causes a system failure that takes several seconds to resolve. Therefore, total loss of communication with the operator represents the most critical of the considered failure scenarios and is assigned the highest priority, resulting in an immediate safe stop. The system behaviour in the event of such a disturbance is illustrated in Figure 7. Unity detects the connection loss to the HoloLens at $t_{\text{WlanInterrupt,Unity}} = 1.842 \pm 0.670$ ms. The MCU sets its digital outputs to switch relays for safe stop at $t_{\text{WlanInterrupt,MCU}} = 40.425 \pm 14.604$ ms, whereas:

- $T_{\text{Unity} \rightarrow \text{ROS}} = 29.610 \pm 14.500$ ms
- $T_{\text{ROS} \rightarrow \text{MCU}} = 8.974 \pm 2.178$ ms

Consistent with the previous measurements, the interface to Unity represents the largest source of varying latency within the system. Due to the rapid detection of the connection loss within Unity, the time interval between communication failure and initiation of the safe stop is comparable to the latency associated with the execution of an active control command by the operator. Consequently, the resulting system response is comparable to that of a stop command intentionally issued by the operator and therefore remains within the system requirements defined in this paper.

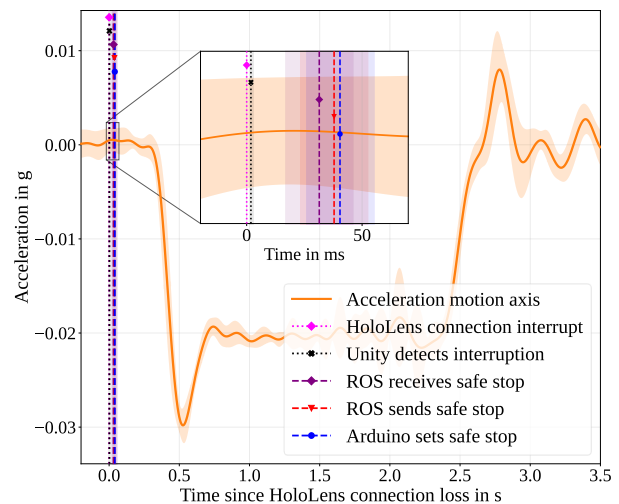


Figure 7. System reaction on HoloLens connection loss

TCP/ROS interrupt. Transient communication interruptions or stalled processes may occur in distributed systems and do not necessarily constitute an immediate safety hazard. To address the effects of such disturbances, watchdog timers are implemented both within the ROS network and on the MCU. A safe stop is triggered only if the predefined timeout period is exceeded. Figure 8 shows the watchdog behaviour of the ROS network on the CCU. Communication with Unity is disrupted by an interrupted TCP bridge, resulting in ROS no longer receiving heartbeat signals. The latency introduced by the AR interface is also evident in the watchdog evaluation results. Although the heartbeat transmission rate is nominally 100 Hz, the effective rate is limited by Unity's internal maximum frame rate. Consequently, the most recently received heartbeat signal is already delayed, causing the watchdog to trigger on average after $t_{\text{TCPInterrupt,ROS}} = 225.937 \pm 14.921$ ms, measured from the moment of signal loss. At the time of watchdog activation, the last valid heartbeat signal from Unity has been 28.224 ± 14.344 ms in average. Accordingly, the ROS watchdog timer reliably triggered after an interruption duration of 254.161 ± 2.198 ms. Consistent with the previous investigations, the processing of the timeout event and the subsequent switching of the Arduino output pins requires low computing costs:

- $T_{\text{ROS} \rightarrow \text{MCU}} = 4.444 \pm 0.665$ ms

The results confirm the robustness of the ROS watchdog in reliably detecting communication failures and triggering the appropriate system response. The low processing latency between CCU and MCU further demonstrates that safety mechanisms implemented at this level can intervene within the low-millisecond range. Such responsiveness is a key requirement for automated controller-based systems.

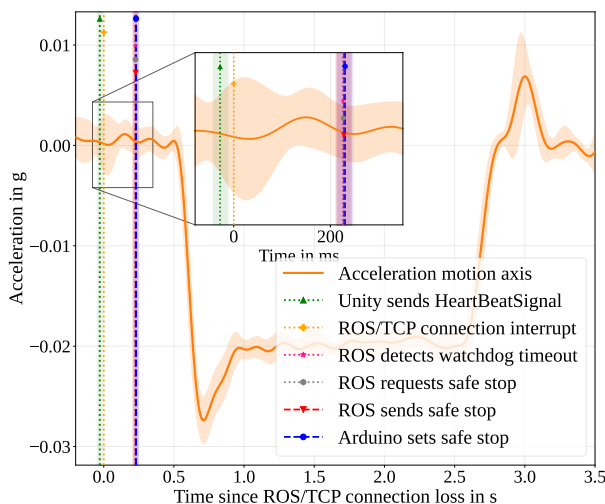


Figure 8. System reaction to TCP failure triggered by ROS watchdog

Serial interrupt. The analysis of the serial watchdog timer is shown in Figure 9. After the serial connection between CCU and MCU is disconnected, the Arduino MEGA entered the safety stop with an average value of $t_{\text{Watchdog,MCU}} = 250.075 \pm 0.004$ ms. As observed for the ROS watchdog, the safe stop is initiated slightly before the configured watchdog timeout elapses, since the last valid heartbeat signal from the ROS node is received, on average, 6.543 ms prior to watchdog activation. This behaviour is expected with the implemented monitoring frequency and serves to validate the low-latency communication with high clock speed between MCU and CCU.

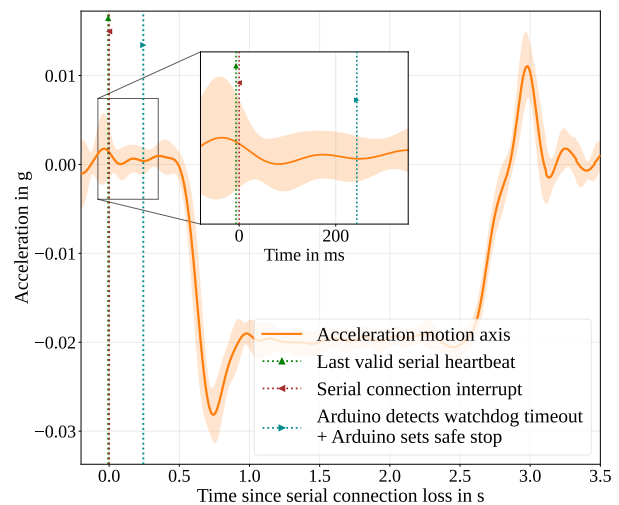


Figure 9. System reaction to serial connection failure triggered by Arduino watchdog

4.4 DISCUSSION

The investigations demonstrate that the implemented HMI reliably transmits operator commands to the crane control system via the ROS-based communication architecture. While all examined communication stages exhibit stable behaviour, the Unity instance introduces the highest degree of uncertainty. The measured standard deviations are significantly larger than those of the remaining system components, as communication performance is directly influenced by Unity's frame rate. Consequently, fluctuations in rendering performance affect the effective transmission frequency, which in some cases is reduced to approximately 12.5 Hz. The Unity interface therefore represents the dominant source of latency and timing uncertainty within the system. However, with regard to the transmission requirements for control signals identified in the literature [15], the measured disturbances do not represent a significant reduction in quality.

In contrast, communication within the ROS network proves to be reliable and exhibits only minor delays. The low communication latency and predictable behaviour of the distributed architecture indicate that the system can be

extended with additional functionalities and provide a suitable basis for future automation approaches. From a safety perspective, the watchdog timers implemented within the ROS network and on the control MCU operate robustly and consistently respond within the predefined watchdog timeout of 250 ms. The fault-injection experiments confirm the reliable detection of communication failures and the correct execution of the corresponding safe-stop procedures.

Despite these positive results, the current implementation does not yet cover all potential failure modes. In particular, a malfunction or deadlock of the MCU itself could prevent the execution of safety functions. Therefore, future work should incorporate an independent hardware watchdog to monitor the operational state of the MCU and enforce a safe stop in the event of a controller failure, thereby further improving the safety and fault tolerance of the overall system.

5 CONCLUSION AND OUTLOOK

This study investigates the response time and failure behaviour of an AR-based control interface for an overhead crane. A demonstrator is implemented comprising an HMI (HoloLens 2 and Unity), a communication layer (ROS2), MCU-based crane interface and an IMU for monitoring the crane behaviour. This allows to evaluate command-processing latency and response to communication failures in repeated experiments.

The results show an average latency of $T_{\text{Unity} \rightarrow \text{ROS}} = 26.120 \pm 16.833$ ms between Unity and the ROS network as well as $T_{\text{ROS} \rightarrow \text{MCU}} = 7.244 \pm 2.686$ ms between ROS and MCU. The total response time from operator input to measurable crane motion is $T_{\text{total}} = 392.260 \pm 45.701$ ms for starting and $T_{\text{total}} = 319.508 \pm 35.579$ ms for stopping. These values are dominated by the response time of the used crane actuating system, showing that the additional latency introduced by the AR system is comparatively small.

To evaluate system safety, various communication failure scenarios are investigated, including a loss of connection to the HoloLens due to wireless network failure, an interrupted TCP bridge and a disconnected CCU-MCU connection. The safe stop commands are executed as specified by the watchdog timer requirements. These findings demonstrate the robustness of the monitoring. Nevertheless, a safe stop regarding to DIN EN ISO 13850 needs to be implemented.

Future work should address the integration of an independent, hardware-based emergency-stop. User studies involving experienced and inexperienced crane operators should assess gesture reliability, usability, cognitive workload and situational awareness. Finally, the architecture can

be extended with environmental sensor fusion, route planning, collision avoidance and eventually allow for fully automated transport functions.

6 FUNDING

This publication was produced as part of IGF Project 01IF23678N (KraNavi) of the Intralogistics/Material Handling and Logistics Systems Research Association (IFL) e. V. through the German Aerospace Center (DLR) as part of the Program for the Promotion of Industrial Collaborative Research (IGF) by the German Federal Ministry for Economic Affairs and Climate Action (BMWE) pursuant to a resolution of the German Bundestag.

7 LIST OF ABBREVIATIONS

AR	Augmented Reality
CCU	Central Computing Unit
HMI	Human Machine Interface
IMU	Inertial Measurement Unit
MCU	Microcontroller Unit
ROS	Robot Operating System
TCP	Transmission Control Protocol
VR	Virtual Reality
XR	Extended Reality

LITERATURE

- [1] W. Xiang, K. Yu, F. Han, L. Fang, D. He, und Q.-L. Han, „Advanced Manufacturing in Industry 5.0: A Survey of Key Enabling Technologies and Future Trends“, *IEEE Trans. Ind. Inform.*, Bd. 20, Nr. 2, S. 1055–1068, Feb. 2024, doi: 10.1109/TII.2023.3274224.
- [2] Deutsche Gesetzliche Unfallversicherung e.V. (DGUV), „Arbeitsunfallgeschehen DGUV 2024“, Okt. 2025.
- [3] Y.-C. Chen, H.-L. Chi, S.-C. Kang, und S.-H. Hsieh, „Attention-Based User Interface Design for a Tele-Operated Crane“, *J. Comput. Civ. Eng.*, Bd. 30, Nr. 3, S. 04015030, Mai 2016, doi: 10.1061/(ASCE)CP.1943-5487.0000489.
- [4] A. Ajikumar und M. Moghaddam, „IoT-Integrated Augmented Reality for Spatially-Aware Human-Machine Interaction“, 6. September 2024. doi: 10.36227/techrxiv.172565688.85662261/v1.
- [5] T. A. Sitompul, „Human–Machine Interface for Remote Crane Operation: A Review“, *Multimodal Technol. Interact.*, Bd. 6, Nr. 6, S. 45, Juni 2022, doi: 10.3390/mti6060045.

- [6] K. C. C. Peng, W. Singhose, und P. Bhaumik, „Using Machine Vision and Hand-Motion Control to Improve Crane Operator Performance“, *IEEE Trans. Syst. Man Cybern. - Part Syst. Hum.*, Bd. 42, Nr. 6, S. 1496–1503, Nov. 2012, doi: 10.1109/TSMCA.2012.2199301.
- [7] A. Duz u. a., „From construction machines to remote construction robots: control, interfaces, and usability of the Cranebot“, *Front. Robot. AI*, Bd. 11, S. 1504317, Jan. 2025, doi: 10.3389/frobt.2024.1504317.
- [8] H. Osumi, M. Kubo, S. Yano, und K. Saito, „Development of tele-operation system for a crane without overshoot in positioning“, in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei: IEEE, Okt. 2010, S. 5799–5805. doi: 10.1109/IROS.2010.5652897.
- [9] C. Yang, H. Yu, Y. Zheng, L. Feng, R. Ala-Laurinaho, und K. Tammi, „A digital twin-driven industrial context-aware system: A case study of overhead crane operation“, *J. Manuf. Syst.*, Bd. 78, S. 394–409, Feb. 2025, doi: 10.1016/j.jmsy.2024.12.006.
- [10] X. Tu, J. Autiosalo, A. Jadid, K. Tammi, und G. Klinker, „A Mixed Reality Interface for a Digital Twin Based Crane“, *Appl. Sci.*, Bd. 11, Nr. 20, S. 9480, Okt. 2021, doi: 10.3390/app11209480.
- [11] F. He, S. K. Ong, und A. Y. C. Nee, „An Integrated Mobile Augmented Reality Digital Twin Monitoring System“, *Computers*, Bd. 10, Nr. 8, S. 99, Aug. 2021, doi: 10.3390/computers10080099.
- [12] S. M. Hasan, K. Lee, D. Moon, S. Kwon, S. Jinwoo, und S. Lee, „Augmented reality and digital twin system for interaction with construction machinery“, *J. Asian Archit. Build. Eng.*, Bd. 21, Nr. 2, S. 564–574, März 2022, doi: 10.1080/13467581.2020.1869557.
- [13] K. Brunnström, E. Dima, T. Qureshi, M. Johanson, M. Andersson, und M. Sjöström, „Latency impact on Quality of Experience in a virtual reality simulator for remote control of machines“, *Signal Process. Image Commun.*, Bd. 89, S. 116005, Nov. 2020, doi: 10.1016/j.image.2020.116005.
- [14] M. Rossander, T. Semberg, M. Englund, und L. Eliasson, „Teleoperated forwarder with VR: impact of latency on operator log handling performance“, *Int. J. For. Eng.*, Bd. 36, Nr. 1, S. 76–88, Jan. 2025, doi: 10.1080/14942119.2024.2415835.
- [15] T. Hatano, T. Shimada, und T. Yoshida, „Evaluation of end-to-end delay requirement for remote, precise pick-up action of miniature crane“, *IEICE Commun. Express*, Bd. 12, Nr. 5, S. 242–248, Mai 2023, doi: 10.1587/comex.2023XBL0017.
- [16] DIN EN IEC 61508: Funktionale Sicherheit sicherheitsbezogener elektrischer/elektronischer/programmierbarer elektronischer Systeme, 2026.
- [17] D. L. Woods, J. M. Wyma, E. W. Yund, T. J. Herron, und B. Reed, „Factors influencing the latency of simple reaction time“, *Front. Hum. Neurosci.*, Bd. 9, März 2015, doi: 10.3389/fnhum.2015.00131.
- [18] DGUV Vorschrift 52: Krane - § 11 Sicherheitsabstände.
- [19] F. Brizzi, L. Peppoloni, A. Graziano, E. D. Stefano, C. A. Avizzano, und E. Ruffaldi, „Effects of Augmented Reality on the Performance of Teleoperated Industrial Assembly Tasks in a Robotic Embodiment“, *IEEE Trans. Hum.-Mach. Syst.*, Bd. 48, Nr. 2, S. 197–206, Apr. 2018, doi: 10.1109/THMS.2017.2782490.
- [20] DIN EN ISO 13850: Sicherheit von Maschinen - Not-Halt-Funktion - Gestaltungsleitsätze, 2016.

Anna-Lena Fritze, M.Sc., Research Assistant at the Institute of Transport and Automation Technology, Leibniz University Hannover.

Anna-Lena Fritze was born 1997 in Langenhagen. She studied Industrial Engineering at Leibniz University Hannover from 2016 to 2023.

Address: Institute of Transport and Automation Technology, Leibniz University Hannover, An der Universität 2, 30812 Garbsen, Germany, Phone: +49 1523762-0533, E-Mail: annalena.fritze@ita.uni-hannover.de

Eike Lorenz, M.Sc., Research Assistant and Project Engineer at Institut für Integrierte Produktion Hannover.

Eike Lorenz was born 1993 in Hamburg, Germany. Between 2013 and 2022 he studied Aerospace Engineering at the TU Braunschweig.

Address: Institut für Integrierte Produktion Hannover gGmbH, Hollerithallee 6, 30419 Hannover, Germany, Phone: +49 511 279 76-234, E-Mail: lorenz@iph-hannover.de

Carsten Schmidt, M.Sc., Research Assistant at the Institute of Transport and Automation Technology, Leibniz University Hannover.

Carsten Schmidt was born 1993 in Willich, Germany. Between 2012 and 2018 he studied Mechanical Engineering at Rhine-Waal University of Applied Sciences, Kleve and Leibniz University Hannover.

Address: Institute of Transport and Automation Technology, Leibniz University Hannover, An der Universität 2, 30812 Garbsen, Germany, Phone: +49 1523762-0542, E-Mail: carsten.schmidt@ita.uni-hannover.de

Dr. -Ing. Benjamin Küster, Head of Production Automation Department at Institut für Integrierte Produktion Hannover.

Dr.-Ing. Benjamin Küster was born in 1988. He studied industrial engineering at Leibniz Universität Hannover. 2020 he finished his doctorate at the Institut für Integrierte Produktion Hannover (IPH).

Address: Institut für Integrierte Produktion Hannover gGmbH, Hollerithallee 6, 30419 Hannover, Germany, Phone: +49 511 279 76-220, E-Mail: kuester@iph-hannover.de

Prof. Dr.-Ing. Ludger Overmeyer, Head of the Institute of Transport and Automation Technology, Leibniz University Hannover.

Prof. Dr.-Ing. Ludger Overmeyer, Head of the Institute of Transport and Automation Technology of Leibniz Universität Hannover. Ludger Overmeyer was born 1964 in Recke, Germany. Between 1984 and 1991, he studied electrical engineering at the University of Hannover. In 1996 he finished his doctorate in mechanical engineering at the University of Hannover. From 1997 to 2001 he worked as project manager, division manager and head of research and development at Mühlbauer AG in Roding. Since 2001 Ludger Overmeyer is Professor at the Institute of Transport and Automation Technology of Leibniz Universität Hannover.

Address: Institut für Transport- und Automatisierungstechnik, Leibniz Universität Hannover, An der Universität 2, 30812 Garbsen, Germany, Phone: +49 511 762-3524, E-Mail: ludger.overmeyer@ita.uni-hannover.de