DOCTORAL THESIS

Extended Reality for I5.0: Towards Human Centricity in Human-Robot Interaction

Simone Luca Pizzagalli

TALLINNA TEHNIKAÜLIKOOL
TALLINN UNIVERSITY OF TECHNOLOGY
TALLINN 2024

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
42/2024

Extended Reality for I5.0: Towards Human Centricity in Human-Robot Interaction

SIMONE LUCA PIZZAGALLI

TALLINN UNIVERSITY OF TECHNOLOGY
School of Engineering
Department of Mechanical and Industrial Engineering

The dissertation was accepted for the defence of the degree of Doctor of Philosophy on 27 May 2024

Supervisor: Prof. Tauno Otto,
Department of Mechanical and Industrial Engineering, School of Engineering,
Tallinn University of Technology
Tallinn, Estonia

Co-supervisor: Prof. Vladimir Kuts,
Department of Mechanical and Industrial Engineering, School of Engineering,
Tallinn University of Technology
Tallinn, Estonia

Opponents: Dr. Leif Oppermann,
Mixed and Augmented Reality Solutions,
Fraunhofer FIT
Sankt Augustin, Germany

Prof. Filippo Sanfilippo,
Department of Engineering Sciences, Faculty of Engineering and Science,
University of Agder (UiA)
Grimstad, Norway

Defence of the thesis: 29 August 2024, Tallinn

Declaration:
Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology, has not been submitted for any academic degree elsewhere.

Simone Luca Pizzagalli ______________________
signature

ISSN 2585-6898 (publication)
ISBN 978-9916-80-178-9 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9916-80-179-6 (PDF)
DOI https://doi.org/10.23658/taltech.42/2024
Printed by Koopia Niini & Rauam

Pizzagalli, S. L. (2024). *Extended Reality for I5.0: Towards Human Centricity in Human-Robot Interaction* [TalTech Press]. https://doi.org/10.23658/taltech.42/2024

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
42/2024

Laiendatud reaalsus Tööstus 5.0 jaoks: inimesekeskse lähenemiseni inimese-roboti suhtluses

SIMONE LUCA PIZZAGALLI

KIRJASTUS

"To Pietro and Matilda"

"The best proof of the specificity of the book is that it is at once a reality of the virtual and a virtuality of the real."

Gaston Bachelard, *The poetics of rêverie*, 1960

Contents

List of publications

The present Ph.D. thesis is based on the following publications that are referred to in the text by Roman numbers.

I **S.L. Pizzagalli**, V. Kuts, and T. Otto. User-centred design in industrial collaborative automated systems. *Proceedings of the Estonian Academy of Sciences*, 70(4):436–443, 2021

II **S.L. Pizzagalli**, V. Kuts, Y. Bondarenko, and T. Otto. Evaluation of virtual reality interface interaction methods for digital twin industrial robot programming and control, a pilot study. In *ASME International Mechanical Engineering Congress and Exposition*, volume 85567, page V02BT02A005. American Society of Mechanical Engineers, 2021

III V. Kuts, J. A. Marvel, M. Aksu, **S.L. Pizzagalli**, M. Sarkans, Y. Bondarenko, and T. Otto. Digital twin as industrial robots manipulation validation tool. *Robotics*, 11(5):113, 2022

IV **S.L. Pizzagalli**, Y. Bondarenko, V. Kuts, E. O'Connell, N. Murray, and T. Otto. Ros-based augmented and virtual reality path planning interface for industrial robotic arms: a preliminary assessment. In *ACM, the 6th International Conference on Robot Systems and Applications (ICRSA), 2023*. ACM, 2024

V **S.L. Pizzagalli**, Y. Bondarenko, B. C. Baykara, A. Niidas, V. Kuts, M. Kerm, and T. Otto. Forestry crane immersive user interface for control and teleoperation. In *ASME International Mechanical Engineering Congress and Exposition*, volume 86649, page V02BT02A021. American Society of Mechanical Engineers, 2022

VI K. Mahmood, **S.L. Pizzagalli**, T. Otto, and I. Symotiuk. Development of an AR-based application for assembly assistance and servicing. In *Procedia CIRP, 34th CIRP design Conference, 3-5 June 2024, Cranfield University, UK*. CIRP, 2024

Author’s contributions

I The Author conducted the research, prepared the figures, and wrote the manuscript.

II The Author developed part of the interface interaction UI, conducted the experiments, analysed the results, prepared the figures, and wrote the manuscript.

III The Author contributed to the manuscript writing and figure preparation, conducted half of the experiments and contributed to the data analysis.

IV The Author developed the interface interaction UI, analysed the results and wrote the manuscript.

V The Author developed part of the interface interaction UI, contributed in the figure preparation and manuscript writing.

VI The Author contributed in the development of the interface interaction UI, prepared the figures, conducted the experiments, analysed the results, and wrote the manuscript.

Abbreviations

API	Application Programming Interface
AR	Augmented Reality
CPS	Cyber-Physical Systems
DOF	Degrees Of Freedom
DT	Digital Twin
HCI	Human Computer Interaction
HMI	Human Machine Interface
HRI	Human Robot Interaction
HRC	Human Robot Collaboration
I4.0.	Industry4.0
I5.0	Industry5.0
IoT	Internet of Things
JSON	JavaScript Object Notation
KPI	Key Performance Parameter
MRTK	Mixed Reality Toolkit
Nasa TLX	Nasa Task Load Index
NUI	Natural User Interface
O4.0	Operator4.0
PLM	Product Lifecycle Management
ROS	Robot Operating System
ROS-I	ROS Industrial standard
SDK	Software Development Kit
SME	Small Medium Enterprise
SUS	System Usability Scale
UI	User Interface
UCD	User Centered Design
UEQ	User Experience Questionnaire
URDF	Unified Robot Description Format
VR	Virtual Reality
XR	Extended Reality

1 Introduction and literature review

1.1 From I4.0 to O4.0 and I5.0

It has been several years since the first definition of Industry4.0 (I4.0) was given [61] and adoption in applied research [110]. After decades of heavy-duty mechanization of production processes lead by hand-based labor, computers made their appearance in the industrial field at some point of the 20th century. The introduction of programmable controllers, electronics, and information technology solutions had a huge impact on both production, overall product quality, market demand, and, most importantly, on operators and their interactions with machines and industrial processes. The initial driving need of I4.0 revolution have been the one for product customization. Consequently the shift, from the old mass production paradigm pushed towards the necessity of flexible manufacturing, shorter product development times, an increased need for decentralized and more sustainable and efficient systems.

We could argue that, other than automation, the other common factor at the base of the so-called pillars of I4.0 is data. Data constitute the driving motor for the interaction of digital and physical world in cyber-physical systems (CPS) [104]. Internet of Things are the source of the physical world state of robots and automated processes; 3D models and advanced simulation the digital counterpart of their physical presence, with additive manufacturing closing the loop between digital models and production of real-world parts and components. Cloud computing constitutes the virtual space for the elaboration of Big Data through advanced machine learning methods. Interoperability and cyber security allow for a safe integration of system data flow. Advanced simulations and Digital Twins (DT) are, in this context, unifying these aspects to provide the well-known synchronized loop between real and digital world [154]. The main goal is the creation of smart factories capable of improve efficiency, reduce errors, optimize the production processes towards and increased flexibility and customization. This is possible thanks to the developed methods and techniques based on the above-mentioned technologies with advanced automation, remote monitoring and control, predictive maintenance, improved training, data driven decision making being the most relevant.

The changes that emerged with increased pace in this domain are related to operators. This has been a permanent characteristic of all the industrial revolutions, including the fourth. The channels, methods and interfaces mediating humans and machine interaction changed drastically with the evolution of industry driving technologies, to a point of modifying, especially in recent times, the skills and roles required from operators. In this context the new industrial paradigm aims at presenting the above mentioned data flow to the end users with the aim of improving their efficiency, safety, productivity on the shop floor. Leveraging human soft skills, decision making, flexibility and intuition, roots on the capabilities of the systems to provide an outline of actionable information thought advanced human machine interfaces (HMIs) [18]. Moreover, the advancement in automated and collaborative robotics shortened the distance of operational space between machines and humans on the physical level, with an increased need for new collaboration methods, metrics and standards [158].

The inclusion of the operator in the information loop together with the evolution of collaborative industrial systems, requires the adoption of novel Human Computer Interaction (HCI) and interfaces and the design of user-friendly UIs allowing the seamless cooperation between machines and humans. Collaborative working setups, on the other hand entail intelligent, adaptable, and aware systems, capable of interacting with, and learning from, humans on a task-based level [14]. This scenario depends upon large investments in the adoption and development of the underlying foundation of integrated and digitalized systems. On top of this an additional effort, in investments and training, must be taken in the direction of visualization technology such head mounted displays (HMD), wearable interaction technologies, and development of applications and interfaces integrating data and systems in different industrial contexts and use cases. This process, especially in small and medium enterprises (SMEs), is still ongoing and largely delayed on many aspects, including old machinery integration, adoption of new collaborative systems, data collection and elaboration, DTs and advanced interfaces adoption [92]. Moreover operators are required to have a higher level of autonomy on the production floor. The level of operators training on the specific technologies effects the support these can give therefore influencing autonomy. This is also dependent on social and economical context. In this respect companies not only face problems on the technical side, but organizational and managerial one [15].

The new role of the operators has direct implications on materials and methods which are more suitable for the purpose, but also human aspects such as health, ethics, data privacy and acceptance of technology, well-being, and safety. A new archetype of Operator4.0 (O4.0) [133] has been proposed to sustain the development from the human perspective and which would be able to address and describe the closer and closer relation between humans, data, simulated environments, and real machines in a sustainable way [134]. O4.0 is now empowered in its cognitive physical and perceptive capabilities by technology in an inclusive and immersive way. The collaboration between CPS and human workforce happens thought the same technologies that lead the fourth industrial revolution but with a renewed central role of the human side which is not only consumer of the information but also producer, and main actor in the loop. Figure 1 exemplifies the changes in human robot interaction within the industrial development from I3.0 to I4.0 collaborative systems towards O4.0 and beyond.

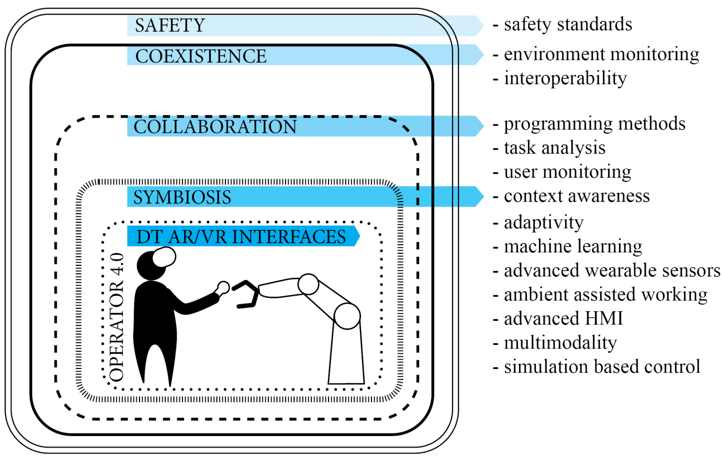

Figure 1: Human Robot Interaction in I4.0 towards I5.0 paradigm. Figure from I.

The figure describes the evolution and layers of human robot interaction and the consequent role of operators in automated robotic system. These changes move from mere safety standards in traditional setups, where man an machine are just coexisting in the same space, to operators being embedded in advanced immersive interfaces. A consequent transformation of technological requirements and support is required moving from environmental standards to intelligent and sensorized systems, closer to the human.

Rapid evolution of technology towards Artificial Intelligence (AI) and cloud computing and immersive DTs in the industrial field, with recent discussions about the Industrial Metaverse [115], together with the newly defined role of humans in the production systems, lead to the revision of the general scope of modern industrial developments. The need to face more consistent and recurrent economical fluctuations and production changes due to disruptive socio-political situations (like COVID-19), oriented the efforts towards new values and objectives in what is now called fifth industrial revolution, Industry 5.0 (I5.0) [31]. The new pillars, or goals, of I5.0 are resiliency, human centricity and sustainability.

- Resiliency refers to the to a higher degree of robustness in value chain solutions, adaptability of production and flexible processes to overcome disruptions. This is also defined by a higher level of customization and new concepts of manufacturing as a service or cloud manufacturing.

- Human centricity, as expected, redefines the role of the operator and humans in relation to technology and machines in industry embracing what already envisioned in [133]. What is clear is that technology is customised and designed to support, guide and train workers while these have an active role in the design and decision making phases. Technology is also fostering well-being, health and safety both in the factory and in general in society.

- Sustainability is referring to the increased need for reuse, recycle and re-purpose while reducing waste emissions and consumption of resources.

The clearer definition of the human role in industry in this respect set in motion a consistent reconsideration of its technological backgrounds, largely derived from I4.0, and how they could match or support the new requirements [109] in different use cases [83]. Humans are envisioned as more and more embedded in the technological framework of productions systems aiming at collaborative symbiotic relationship with tangible and non tangible assets. This requires solutions that would mediate and define this interaction in each point of the production and product life cycle but most importantly with each of the mentioned technologies [161]. It is also clear that both human and machine cognitive aspects are greatly valued in I5.0. The accelerated development and dissemination of AI pushed forward one of the conditions of the paradigm, which is machines and systems supporting in an intelligent, context aware, adaptable way the operators actions and decisions, while these contribute with typically human soft skills, creativity and decision making. The synergy between humans and CPS is central in this scenario. Nevertheless, the integration of operators and technology mediated by cognitive, sensing and interaction systems, raises questions and concerns on ethical, social, psychological and organisational aspects which need to be confronted to grant an efficient and healthy adoption of I5.0 on large scale [82], [37].

1.2 Extended Reality and Digital Twins for Human-Robot Interaction

1.2.1 Augmented and Virtual Reality

The first applications of Augmented and Virtual reality technologies date back to experimental systems in the early 60s. The last sixty year have been studded by examples working in the direction of creating interactive or immersive user experiences such as

Sensorama, patented in 1963, Ultimate Display, [145], Sorcerer's apprentice [156], or Put-That-There [20] to name a few. AR/VR technologies have the potential to accommodate, on different degrees and through different means, the virtual and real world in the same scope, ultimately facilitating the visualization and interaction with both simulated and real world objects and systems. This virtual continuum [26], shown in Figure 2, goes now under the umbrella term of Extended Reality (XR).

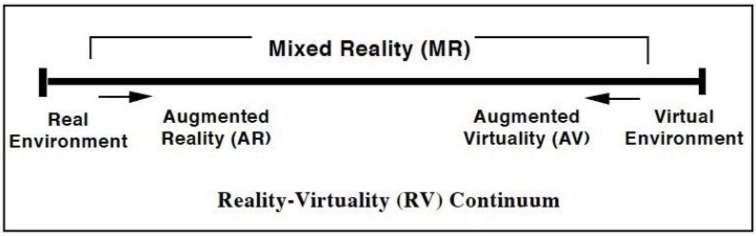

Figure 2: The virtual continuum according to Milgram. Figure from [96]

Moreover, XR can be defined as highly advanced HMI capable of working in quasi real time, providing experiences from different sensorial channels (e.g., auditory, visual) and allowing the user to make changes to the computer generated scenario [26]. This matches very well, incidentally, with the user centered expectations of I5.0 outlined above. In this respect the capabilities of both inputs and outputs given by devices are always in evolution [13] with hardware and software interaction and visualisation methods capable of reducing the gap between the real and virtual realms. This include a vast domain of solutions spanning from the well known HMD for the visualisation, to tracking devices, either mechanical, markerless, based on inertial or magnetic sensors, depth cameras, eye tracking systems, electroencephalography or electromyography signals, forced feedback, data gloves for virtual object interactions, treadmills for continuous walking, and so on.

The efficiency and effectiveness of these systems and methods is highly dependable on the type of use case requirements and bound, in any case, to their impact on the user. For industrial applications, for instance, different evaluation metrics can be applied to define which solution would be the best in every scenario. Field of view, ease of maintenance, mobility, immersiveness level, number of allowed concurrent users, ergonomics, tracking capabilities, and feedback, are among the most common evaluation criteria [136]. This does not include software-based metaphors and illusions which are capable to amplify the interaction possibilities of the available hardware by overlapping fully virtual capabilities with real world controllers (for instance extending a virtual arm to grab an object). Because of the exceptional range of potential applications and goals of the technology not all the above mentioned requirements, despite being crucial, are essential prerequisites for each use case application [19]. Albeit the large adoption and economical forecasts and expectations towards XR [11] there are still a number of open issues that need to be solved. The main open challenges relate to the level of accuracy, cost effectiveness, reliability and data security [129]. Human factors such as ergonomics, usability of virtual and augmented interfaces, information visibility, privacy and user acceptance have a high impact on technology adoption especially in SMEs [42] [91]. Values and benefits of these technologies are general still difficult to communicate [19]. Furthermore, user related physiological issues may arise from the prolonged use of the devices [139] such as, to name the most common, nausea, associated to image fidelity, systems latency and more in general the mismatch between visual and vestibular inputs [30], and eye fatigue, which is related to differences between focus point and vergence due to the closeness of monitors in HMD

[163].

Regardless, XR technologies prove to be valuable tools in several industrial, and non industrial, fields [33] [35]. Example applications, restricted to the manufacturing domain, includes the adoption in the following use cases:

- Improvements of control and safety in human robot collaboration [160] [85] [39].
- In the forestry sector (topic addressed in V) with examples of crane control and teleoperation using immersive UIs [128] [121] [164], planning and simulation [87] [66], training crane operators [122] and trajectory planning [117].
- In applications for robot programming [114] and control, including path planning tasks [153] [119] [126].
- Teleoperation of robotic arms in different use cases, [166] [142] [108] [113] [138] [79].
- As a base for the integration of the operator in the CPS for multi robot control [167].
- As a tool to assess user experience [131] and improve product design [106]
- Operator support and training [40] [97] [157] in maintenance and assembly tasks [44] [141] [77] and which is a topic addressed in VI.

A few examples adopting XR and DTs for HRI together with attempts to define metrics and evaluation procedures are discussed in the following paragraphs. But let's first frame a few more concepts related to this study before comparing a few examples.

1.2.2 Digital Twins

Digital Twins are advanced representations, not necessarily three dimensional, of the status of a product, a complex system or processes defined by a real-time, synchronized bidirectional loop transmission of data between the digital and real entities. The definition given by Grieves [47] has now developed and expanded based on the many characterisations [16] and adoption fields of this pervasive technology. Examples can be found in aviation, health [76] [24], management of large living environments [64], smart cities and mobility [130], building construction [140], and of course industry and manufacturing. The growth of interests and research conducted on the topic has been consistent for decades [80] with real world applications focusing on optimization and scheduling, human robot interaction, system monitoring, product development [81], predictive maintenance and virtual tests [148]. DT have become an industrial standard technology, nevertheless there is a lack of shared validation leading to fragmented implementation in the industrial context and clarity in investment return especially in SMEs [60]. The definition of Key Performance Parameters (KPI) on a medium long term, and the integration with other key technologies such as XR and high speed connectivity, is essential to determine the benefits and return of investment. This topic is partially discussed in VI. Moreover, it is important to evaluate the standards and requirements of industrial IoT technologies, upon which DTs heavily rely, to comprehend their suitability across various use case scenarios. This is also true when the technology is adopted along the whole product life-cycle, and which seems to be still a rare case [80]. The fidelity level of a digital twin seems to hinge on specific application requirements, desired performance levels, and cost-effectiveness, while addressing policies and regulations concerning the vast amount of generated data flow remains a significant concern.

The core of the twinning process lies in the interpretation and description of both physical and virtual entity states, alongside a phase dedicated to data collection and synchronization [60]. The first phase is characterised by data acquisition thought IoT and extended use of sensors, employed for the description of the systems and the containing environment. After processing and interpretation, data is used to change the state of either the physical or digital counterpart. This happens, in the former case, thought the adoption of actuators.

The main conceptual, and practical difference between digital model and digital shadow against DTs belongs to the synchronization phase [70].

- DT comprise of a two-way automatic synchronization between physical and digital asset.
- Digital models are simulations that can run completely independent data analysis on a specific system without being connected to a real world asset. The transmission of data can nevertheless happen in a manual way.
- In digital shadows the connection between the real assets and simulation is automatic and synchronized only in one direction while the feedback from digital to real is missing or on demand.

The context is an important factor in the definition of the parameters involved in the twinning process. These includes, for instance, geometric description, asset functions, machines health state, location, performance, hierarchical relations to other resources. Without entering the details of twinning rate, accuracy and fidelity which depend on different characteristics of the DT and parameters collected and processed during the loop, it is worth mentioning how DT technology is often adopted for machine to machine and human to machine interfaces [155]. In this context a human perspective has been only partially addressed both in terms of human inclusion in the loop and its description [46][25], nor in the definition of requirements and design frameworks for intuitive and accessible DTs UI [16]. This interaction is at the core of this work and will be further discussed in the following paragraphs.

1.2.3 Human-Machine Interfaces and collaborative systems

HCI refers to the manner in which users achieve objectives by interacting with a computer system, and which are responding with suitable feedback and modifications of the world state. This principle applies to various types of interactions between humans and technological artifacts, such as machines and robots. What distinguishes the flow of information between humans and machines is the execution time, during which a user formalizes an objective and takes action within the system, and the evaluation phase, in which the system perceives changes in the world state, interprets them, and communicates the results to the user. Serving as the layer that completes and facilitates this communication loop is the system interface. Interface design, or the design on how human interact with the surrounding world has long history and has been widely formalised in theory and practice [111].

Actions performed on the system can be of different nature depending on the interface readiness for interpretation of specific inputs through several communication channels [59] This can happen thought Natural User Interfaces (NUI) grounding on interaction methods that area already commonly used among humans including gestures and vocal input, for instance. Tangible User Interfaces utilise physical object manipulation to control other physical or digital systems [48]. The use of physiological based signals is also more and more popular, with examples exploiting eye tracking, facial expressions electroencephalographic or electromyographic signals.

Interfaces that are able to transfer signals from the user to the system and back on more than one channel in a coordinated way are called multimodal [41]. Advanced interfaces for HRC should provide valuable information and feedback to the human partner and be able to adapt to the specific context of use or user characteristics in an automatic and semi-automatic way. This characteristic is also called adaptivity of the user interface [63]. Adaptivity consist of some modification of the constituents features of the interface (e.g. interaction modality, output media) driven by some external factors (e.g. type of task, environment, user skills). This adaptation serves various objectives, which may be user-based, aimed at assisting operators with different characteristics in task performance, or system performance-based, such as minimizing errors or enhancing efficiency. Adaptability, unlike profile-based adaptations, involve a learning process grounded in teaching cases, data collection, and analysis. Figure 3 illustrates the HRI loop and clarifies the role of interfaces and interface evaluation.

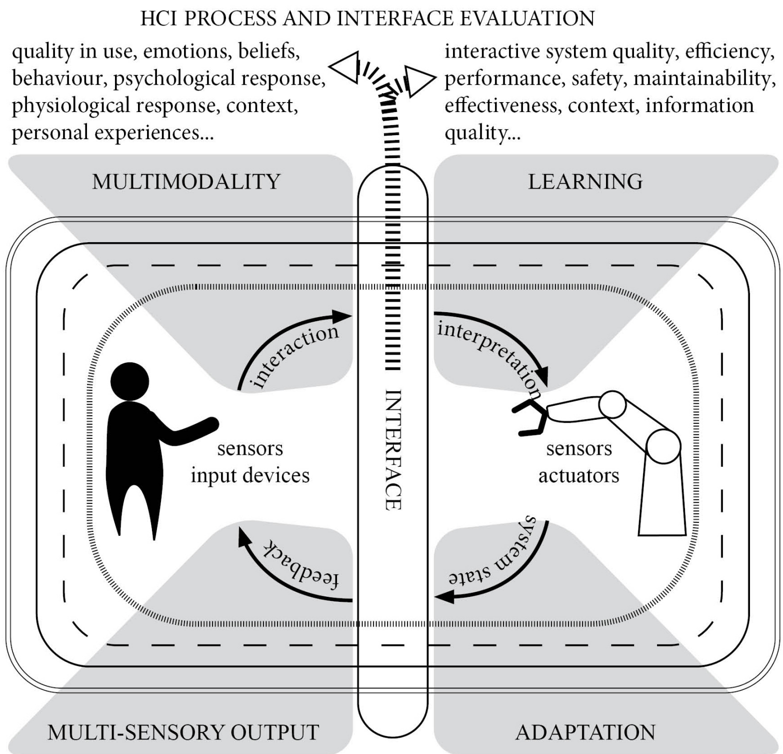

Figure 3: Role of interfaces for HRI in Industry5.0

The described HCI needs to be contextualised within the specific scenario of I5.0 requirements and modern Human Robot Collaboration paradigms (HRC). HRC is defined in ISO [56] and comprises different levels of interaction between the operator and robotic system [51]. These can be described as:

- Independent work, when human and robot perform different tasks on different pieces.

- Synchronized work, when human and robot share a work piece but act on it independently, in different moments in time, in different spaces and by means of different tools.
- Simultaneous work, when the work piece is shared but there is no contact between human and robot.
- Collaborative work, when human and robot work at the same time, on the same work piece while sharing tools and working space.

It is evident, and has already been mentioned, that there are many open challenges for the forthcoming adoption of collaborative systems [158] including, for example, interface design and control, input/output modalities and context awareness. Safety and programming methods are two of the most crucial [34].

Programming methods are of particular interest in the way the procedure is mediated by specific human and machine interactions which are, also in this case, increasingly collaborative and characterised by proximity. [159]. These methods can be distinguished in:

- Offline programming. In this case the program is first created on a separate computer software or simulation, tested and afterwards transferred to the real world robot.
- Online programming. In this case the programming happens directly on the machine through teach pendant and native controllers.
- Walk-thought programming is typical of modern collaborative systems as the operator manipulates the robot while path points or trajectories are automatically saved.
- Learning by Demonstration. In this case the systems learns the required actions and tasks by observing the operator actions and movements [14]. XR is largely adopted for this task as it provides embedded posture and position tracking [36].

As described in Figure 1 the closer the human and machine are, the higher, diversified and complex the systems requirements will be. Evaluation metrics and assessment approaches for both system and humans constitutes a vast domain in HRI, comprising qualitative and quantitative methods and metrics [90] including: information quality and efficacy, communication and time variables, situation awareness, mental effort, physical effort, human response (physiological and psychological), user experience, usability and task performance. Context and agent awareness, system and controls visibility, information availability together with improved usability seem to be the key success elements in advanced human-robot interfaces [28]. Collaboration driven UIs support the way humans and machines communicate with each other in a effortless and efficient way, foster context awareness, and human understandable interactions within the digital-physical world connection loop.

1.2.4 User Centered Design

User Centered Design (UCD) is a well established and standardized method [58] for the design of interactive systems. This approach is by definition multidisciplinary and iterative. The crucial point of this method is the inherent capacity of involving different needs, goals and stakeholders in each of the design process steps. Figure 4 exemplifies these phases and some of each relative requisite definition.

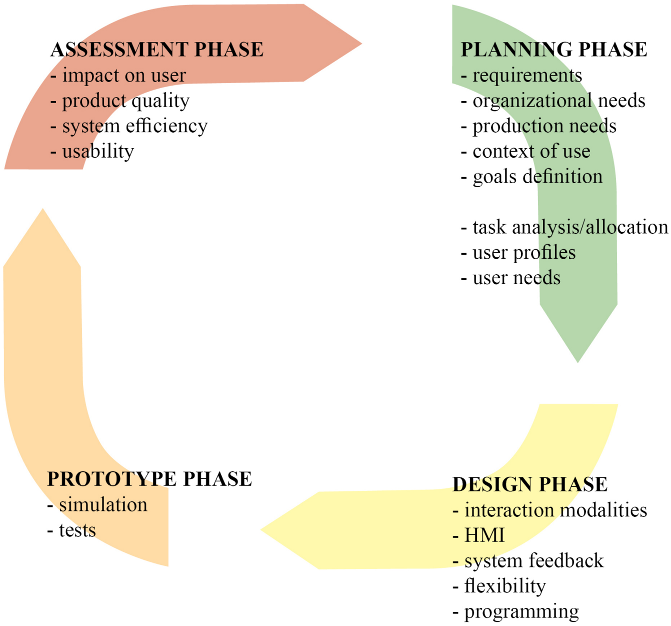

Figure 4: User Centered Design iteration process

The iterative process involves the planning, design, test, and assessment of provided system and interface solutions.

- The planning phase consists in the initial definition of requirements from both production, organisation and stakeholders point of view. The necessary information, in this case, can be gathered thought established methods such as interviews, focus groups, questionnaires, direct observations, task analysis, for the specific use cases, definition of user profiles and personas. Quantitative data sources and real-time simulations can assist in evaluating users', environmental, and organizational aspects by providing insights into the state of the system in all aspects. This has been partially addressed in the use case presented in VI.

- The second phase consists in the design of the interactive systems or interface with related selection of input and output modalities, hardware and software solutions depending on the defined goals and tasks between human and machine.

- The prototyping phase aims testing initial development of the application and allow for a revision of both previous steps.

- The assessment phase takes care of evaluation system quality, efficiency, effectiveness, impact on the involved stakeholders, usability and user experience, based on direct data collection and analysis, and specific use case related KPIs.

The assessment phase is of particular interest in the scope of this work both the technical aspects and user related ones and specifically for what concerns the work done in II, III, and IV. Software quality assessment metrics, for instance, are defined in [55]. This standard divides the evaluation in two main domains: quality in use, and product quality characteristics. The latter are related to technical features of the application, such as, performance, efficiency, compatibility, security, portability etc. The former is related to how the user achieves the goal or satisfies a need through the interaction with the system. The metrics that delineate the success of this process include effectiveness, efficiency, risk mitigation, and satisfaction within a defined context. The concept of usability defined in [57] mostly overlaps with this human centered metrics as it aims at evaluating efficiency, effectiveness and satisfaction in similar goal oriented scenarios. The first two criteria are largely data driven and can be assessed by quantitative methods such as, number of errors, number of successful interactions or task accomplishments, time, resources employed to achieve the goal etc. Satisfaction is linked to the less mathematical quantifiable user preferences, which largely depend on an a set of personal characteristics and that are always different. User Experience (UX) is defined in [58] as "a person's perceptions and responses that result from the use or anticipated use of a product, system or service". This includes quantifiable and other less definable metrics such as emotions, beliefs, preferences, perceptions, responses both of a physical and psychological nature and which are bound to a time frame including moments before and after the actual interaction. Both hedonic and pragmatic qualities of the system need to be detected by means of qualitative and quantitative data collection.

- Quantitative data consist of performance related metrics, such as success rate, number of errors, time to accomplish a specific task, precision which have been utilized in II, III, and IV. Other information can be collected from the user directly by adopting wearable sensors. These methods are largely employed in assessing psychological and physiological changes during the interaction. Heart rate (III) is adopted together with skin conductance response to evaluate the user arousal, and stress levels. Eye tracking is often used to either interact or detect level of attention on some specific system feature (III). Other type of sensors, such as depth cameras, active of passive tracking systems, might give insight on body posture and help estimate ergonomics on the work place.

- Qualitative evaluation methods include interviews, focus groups, speak aloud protocol, observations of the user while using the system and performing a specific task, and use of structured and semi structured questionnaires. These are also adopted in the evaluations discussed in II, III, IV, and VI.

- There are also many subjective quantitative evaluation metrics based on validated questionnaires which are used to detect different aspects of the user experience. NASA Task Load Index (Nasa TLX) [49] (II, III, IV, and VI) aims at assessing the task load and impact on the user including mental and physical effort while performing a task. The User Experience Questionnaire (UEQ) [78] assesses both hedonic and pragmatic qualities related to user experience. System Usability Scale (SUS) [23] attempts giving an overview of subjective evaluation of usability of a system II, IV, and VI. Godspeed questionnaire [17] (used in III) aims at measuring the users perception of robotic systems.

Many more evaluation methods, tools and metrics exists for different aspects of the interaction with both XR systems, physical work spaces and, already mentioned, HRI [90]. These include, for instance, the measurement of flow [132], sense of presence [137] or simulation sickness [67], in digitally generated environments, or, related to real world workstations the evaluation of upper limbs ergonomics and related factors in repetitive tasks [65] [94].

1.2.5 Digital Twin based Extended Reality applications for Human-Robot Interaction

It is worth giving an overview of the different contributions and attempts to integrate DT and XR technologies in HRI industrial domain related use cases. Table 1 provides an overview of these applications.

Table 1: Comparison of DT based XR interface for DT setup, interaction hardware and software, interaction modality and assessment metrics.

	Simulation and DT	Interaction HW SW	Interaction Modality	Evaluation metrics
Teleoperation				
[76]	VR environment synchronized on a 4G connection and open VPN server	HTC Vive HMD and controllers	Dedicated VR UI and mediated by hardware controllers	Cybersickness (not specified)
[32]	VR environment hosted on a Photon server	HTC Vive HMD and controllers	Mediated by hardware controllers	Heart Rate, Electrodermal activity, Eye-tracking
[166]	3D models synchronized via Vikon motion capture system	HTC Vive HMD, Custom haptic glove	Hand manipulation in VR	-
Safety				
[89]	VR scenario synchronized on Arduino Mega	Keyboard and mouse, Kinect V2	Motion and posture	-
[120]	VR environment based on 3D cad models	HTC Vive HMD and controllers, KinectV2	Mediated by hardware controllers	Acceleratio, kinetic energy, leaning angle, movement direction of operator, force related danger, custom questionnaires
Commissioning				
[124]	VR environment based on photogrammetry and synchronized on a non-specified control system	HTC Vive HMD and monitor-based interface	Mediated by hardware controllers	-
[50]	VR DT based on Catia and Modelica, zero MQ framework-based server	HTC Vive HMD and controllers	Mediated by hardware controllers	Body posture based on Percetion Neuron Pro, RULA Score ergonomics assessment
[86]	VR environment based on Siemens NX and Technomatrix Process Simulate	HTC Vive HMD and controllers, KinectV2	Dedicated VR UI and Mediated by hardware controllers	Collision analysis, reach test, placement test, vision test
[125]	VR environment, Technomatix Jack for operator avatars and Vicon motion capture	HTC Vive HMD and controllers, Oculus Rift, Rear projected monitor, Nintendo Wiimote	Mediated by hardware controllers	Heart Rate, Heart Rate Variability, Breath Rate, Eye gaze and pupil dilation, Postural analysis with RULA score and OCRA index, Heuristic analysis, NASA-TLX
Programming				
[27]	Existing Cad models and Robostudio	Oculus Rift HMD and controllers	Mediated by hardware controllers	-
[144]	Bullet Physics as physics engine simulation and URDF robot model	HTC Vive HMD and controllers	Vocal Input and mediated by hardware controllers	-
[118]	Robot models are taken from KuKa experimental package, HoloLens environment reconstruction	Microsoft HoloLens	Custom made handheld pointed, gestures	Time to complete the task, custom usability questionnaire
[88]	Virtools	Custom made tool, 6 DOF tracking sensor	Manipulation, photogrammetry based tracking	-

Training				
[93]	VR environment based on custom 3D models, Evolver for avatar modelling	eMargin z800 3D visor, Kinect, keyboard and mouse	Gestures	Custom presence by involvement questionnaire
[103]	Camera based object positioning	Keyboard and mouse, Tablet, Oculus rift DK2, Google Cardboard, 3D mouse	Mediated by hardware controllers	Focus of attention based on eye tracking
[123]	VR environment based environment scanning	HTC Vive HMD and controllers, Oculus Rift HMD and controllers	Custom VR UI, Mediated by hardware controllers	Custom interfaces evaluation questionnaire
[162]	Optimised 3D models and camera based acquisition	HTC Vive and controllers	Custom made welding tool based on HTC controller	Welding accuracy data analysis
UI Evaluation				
[135]	Point cloud and URDF robot model	Keyboard and mouse, Hololens, Xbox controller	Mediated by hardware controllers	Task completion time, SUS, Nasa TLX
[165]	Point Cloud and URDF robot model	Keyboard and mouse, Hand Tracking via KinectV2, HTC Vive HMD and controllers	Direct manipulation and mediated by hardware controllers	Task completion time, SUS, Nasa TLX
[71]	3D models synchronized via ROSbridge and local network	Microsoft HoloLens	Gestures, Voice	NASA-TLX, SUS, AttrakDiff usability questionnaire, completion time, accuracy
[146]	VR environment based on aeroengines 3D models and real-time data plots	Oculus Quest, Leap Motion	Gaze Input, gesture, dedicated VR UI	Simulation Sickness (SSQ), Flow Short Scale (FSS), NASA-TLX, Igroup Presence Questionnaire (IPQ)
[53]	Wolfram Mathematica and Linkage Designer	Xbox One and Kinect	Gesture	-
[52]	Point cloud and URDF based	HTC Vive, Kinect	Mediated by hardware controllers	SUS, NasaTLX, time to complete the task, accuracy

Recent studies emphasize the importance of a human-centered approach in collaborative industrial systems. The attempts are based on limited examples focused mainly on XR technologies adoption and validated by comparing previous studies approaches [45]. In general, there is inconsistency in the goals, evaluation methods, metrics, and validation used across various experimental scenarios while no clear focus is given to DT based XR interfaces for HRI as a tool to integrate the user in the loop and assess impact and performance.

2 Research gaps and motivations

From the technologies and methods proposed in the framework of the I4.0, and which are not yet consolidated, neither concerning validation nor large scale adoption, especially in SMEs, the scenario is already drifting towards a different industrial paradigm. I5.0 has its roots in the cyber-physical systems technological developments and know-how but takes it from there to aim at sustainability, resiliency and, most importantly user centricity. This new paradigm provides a vision, while missing an implementation plan that would allow its objectives to become factual. There is a confused approach to multidisciplinary development in different domains related to O4.0 and I5.0, including advanced interfaces, such as XR UIs, collaborative systems developement, well-being, psychological and ergonomic aspects, which need to be allocated in novel design framework. Moreover, there is a fragmentation of materials and methods in the definition of the role of XR interfaces for HRI, lack of structured user studies and validation, especially in real world use cases.

2.1 Motivations

As already mentioned we argue that DT based XR interfaces are and will be more and more central tools in the transition to new collaborative industrial paradigms. Additionally, we believe they can constitute a method, and not just a technological mean, to address user centricity in the context of I5.0. These technologies inherently place end users at the centre of extended visualisation, interaction and control capabilities aligning with the ultimate objectives of Industry 5.0 in redefining the roles of operators. To validate the potential of this technologies an extensive work should be done in the definition of a user centric approach that would unify materials and methods on the technical side, on the design and development process, and on the assessment and validation of these technologies on the user for the specific use cases.

Defining user centricity If the user is and will be more and more central in this development, we support the idea that design, integration, interactions and assessment methods, metrics and tools should be also user centric. There should be a change in focus from machines and DTs design and evaluation processes to the operators, both in their interaction with the enabling hardware and the one with the digital environments. The operator empowerment thought cyber-physical systems should insists on a architecture integrating humans, automated systems, DTs, production, and management needs in the same workflow.

A resilient framework If production processes, physical systems and manufacturing units, with examples of manufacturing as a service, are going to be more flexible, on demand and prone to market fluctuations, the methods that would allow designing and assessing new type of interfaces in the scope of Industry5.0 user centricity goals, should keep up by being agile and adjustable to specific requirements and conditions. As discussed before, current examples attempting a more defined approach to the operator in the loop perspective are prone to fragmentation. This is true for both technical approaches, software and hardware solutions, evaluation methods, metrics, and the pletora of use cases in research and industry that, despite the efforts, don't get to a point of a consistent development moving forward but rather in circles around the problem. The type of technology allowing for XR experiences is in continuous evolution and highly influencing the type of interaction and control methods. A set of shared and consolidated best practices, materials and methods, for design and assessment of DT based XR interfaces would allow applying these in a variety of different use cases without starting from scratch or being too tied to technological advancements.

Adoption of DT based XR technologies These brings us to our next motivation which lies in the need of successfully and efficiently facilitate the adoption XR technologies within the industrial domain especially in SMEs. This process is slowed down by many factors including lack of resources, lack of training and support both on managerial and operators level, limited knowledge of the potential applications and available technologies. The level of autonomy required by the I5.0 is higher, as technology is more and more supportive ad adaptable. But requirements for these are also dependable from social and cultural aspects. There is a need of testing methods and materials on a larger samples with different education and cultural backgrounds. In general, an approach that would consider needs and requirements on different levels could facilitate a more efficient decision making flow, towards adoption of advanced XR interfaces in different real world use cases and user groups.

Developing and validating DT based XR interfaces in HRI At the best of our knowledge the prevailing design and assessment methods are not yet sufficient to provide a resilient plan that would include human factors as central pivot in design and decision-making scenarios related HRI and I5.0 technology adoption. We aim at developing and validating DT based XR interfaces based on the proposed approach for human robot interaction systems.

2.2 Research objectives

The main research objective is to develop a dedicated framework for HRI supporting the UCD and I5.0 principles, in particular a user centric approach, through the adoption of DT based XR interfaces as a core method.

The overall research objectives of this work can be summarised as follow:

RO1 It is evident that new design and assessment methods are paramount to favour a transition to new production requirements and scenarios. The first research objective aims at introducing a design approach and test a specific set of tools for DT based XR interfaces validation in supporting human centricity in modern production systems, especially in SMEs, towards I5.0 HRI and collaboration methods with different robotic cells.

RO2 Another objective of this work is transferring the focus of the evaluation from the DT and physical machines to the operators, both in the interaction with the hardware and within the Virtual Environments while preliminarily prove the efficiency, effectiveness and satisfaction of DT based XR interfaces, against native controls for specific use cases. A focus is given to HRI with different robotic cells and operators support on the production floor. Insights on the impacts of interaction practices would provide guidance for industrial developers when determining future factory assignments adopting the specific technologies.

RO3 Another goal is applying the evaluation framework and testing different assessment methods for DT based XR interfaces and interactions, by means of user studies, draw some preliminary conclusions to improve metrics to be applied for the specific use cases.

RO4 Last but not least the final objective of this research is to take the proposed methods, tools and assessment procedures to real world use cases within local SMEs.

2.3 Research questions

RQ1 What role do DT-based XR interfaces play in developing a user-centered approach for HRC systems in Industry 5.0, and how well do they meet the new industrial requirements for safer, collaborative, resilient, and user-centric systems?

RQ2 What is the reliability of DT XR interfaces compared to traditional controls in specific use cases and how do DT-based XR tools perform in terms of effectiveness, efficiency and satisfaction compared to traditional technology counterparts?

RQ3 Which are reliable tools and metrics to assess effectiveness and efficiency of DT based XR interfaces in HRI for I5.0 solutions?

RQ4 Can the developed framework of methods and tools be scaled and validated in real world use cases for SMEs transitioning to I4.0 and I 5.0 solutions?

2.4 Scientific contribution

The presented research questions are answered in the provided scientific contribution as follows.

- RQ1 is answered through the development of DT based XR interfaces and their validation thought a User-Centered Design approach with improvements and testing of the selected technologies and methods along different applications with lab experiments and on field trials.
- RQ2 is answered through design experiments, user studies and lab based user evaluation of the proposed use cases for HRI (over 100 subjects overall).
- RQ3 is achieved, at least partially, thought the integration of different tools and the validation of specific metrics. To answer this questions, different qualitative and quantitative data analysis are adopted in different laboratory experimental sessions (task data collection, wearable sensors for bio-metric data, eye tracking, validated questionnaires for user experience, usability, task load, godspeed etc.)
- RQ4 is answered, in some measure, through the integration of the proposed UI tools and design methods on two real world use case scenarios.

Table 2 summarizes the scientific contribution for the proposed research questions.

Table 2: Relationship between the research questions and the included papers

RQ	Paper I	Paper II	Paper III	Paper IV	Paper V	Paper VI
RQ1	✓	✓	✓	✓	✓	✓
RQ2		✓	✓	✓		
RQ3		✓	✓	✓		
RQ4					✓	✓

3 Towards human centricity in advanced collaborative manufacturing and HRI (I)

Based on the topics discussed in section 2 a programmatic research and development framework is proposed. The idea aims at integrating the previously discussed 'loops' into a unified design and evaluation process rooted on the adoption of DT XR user interfaces.

Initially, we aim at clarifying the synergies between the HRC layers of interactions and related technologies, with the UCD loop (Figures 1 and 4) and summarised in Figure 5.

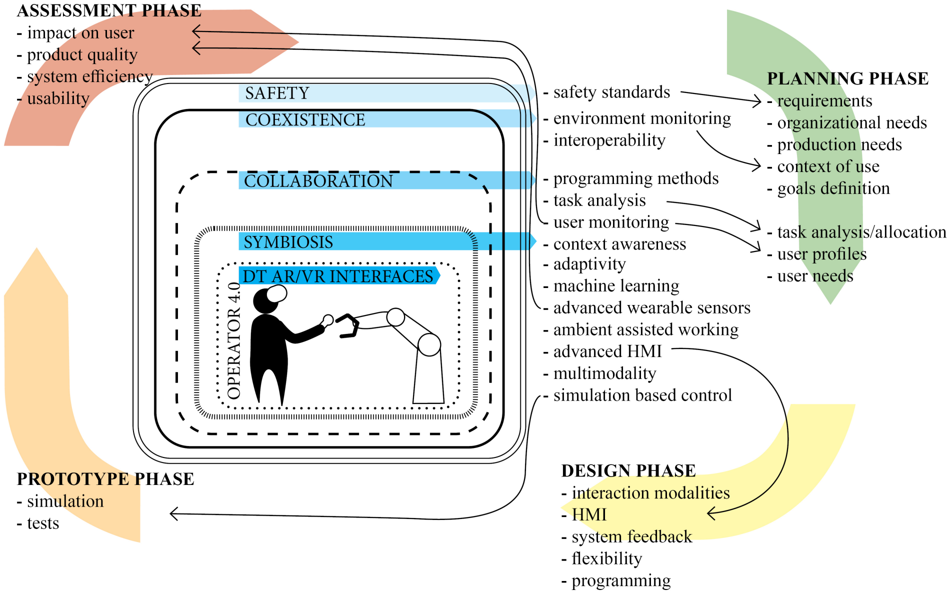

Figure 5: Integration of UCD approach and HRI collaboration layers

Requirements, requirement definition methods and development technologies in HRI are moving toward the operator as a central resource in the control and management of complex, supposedly smart, collaborative systems. Moreover, technological solutions are developed in the perspective of operator support and empowerment. These are largely matching in both UCD and HRC domains. As already mentioned, this brings the focus on the definition of what are the most appropriate technologies granting an efficient, effective yet sustainable interactions between human and machines. XR can be the core solution for many already discussed reasons including:

- It allows for user integration and monitoring through inherent hardware solutions.
- XR equipment is natively multichannel, supporting many different interaction and system feedback (visual, auditory, vocal, haptic, gesture and posture based).
- It allows for online and offline programming, testing, data visualization and teleportation of different machines through direct interaction with their DTs.
- It favours training, and operator support in different tasks (assembly, monitoring, control and teleoperation).

- It supports iterative testing and assessment in a protected environment with repeatable conditions [62]. This is also one of the motivations of the wide adoption of this technology in medical and rehabilitation field [100],[98].

The proposed user-centric scenario grounds, in an initial stage, on existing developed DT based software integration [74], Figure 6. The goal is to further improve XR UIs, interaction methods and test them against native machine interfaces and controllers following a iterative UCD approach.

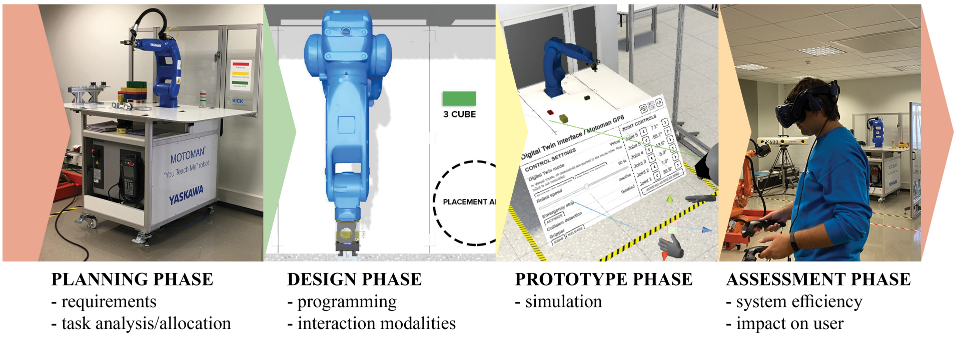

Figure 6: Use cases for DT based XR development and assessment

Figure 7 attempts a systematization of the requirements for each software package involved in the development. This is not a programmatic representation as much as an attempt to organise what discussed in previous sections in terms of immersive HMI, context and machine features, twinning and synchronization, characterisation of user and HCI strategies.

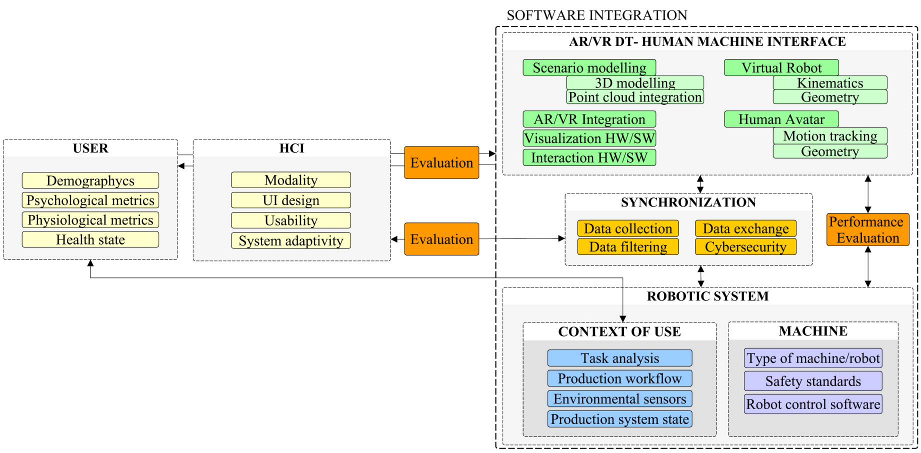

Figure 7: Software integration architecture for UCD based HRC systems adopting DT XR interfaces.

This type of integration needs a standardisation of software packages and testing methodologies allowing homogeneous experimentation and repeatability on similar use cases, but different contexts. There are recent examples going in this directions and creating community based research strategies to improve and accelerate validation and testing of

XR technologies especially in relation to user studies, and validation with target population [9]. In this respect we propose a programmatic, and pragmatic, approach by referring to the presented integration of HRC and HMI diagram demonstrated in Figure 3 and expanding the interface layer to disclose the potential, and limitations, of XR as an high-end interface. Figure 8 illustrates the HCI loop happening both within the DT based XR interfaces between the human avatar and DTs, and without, with the devices granting access to the DT loop on the user end, and to the robotic equipment on the other.

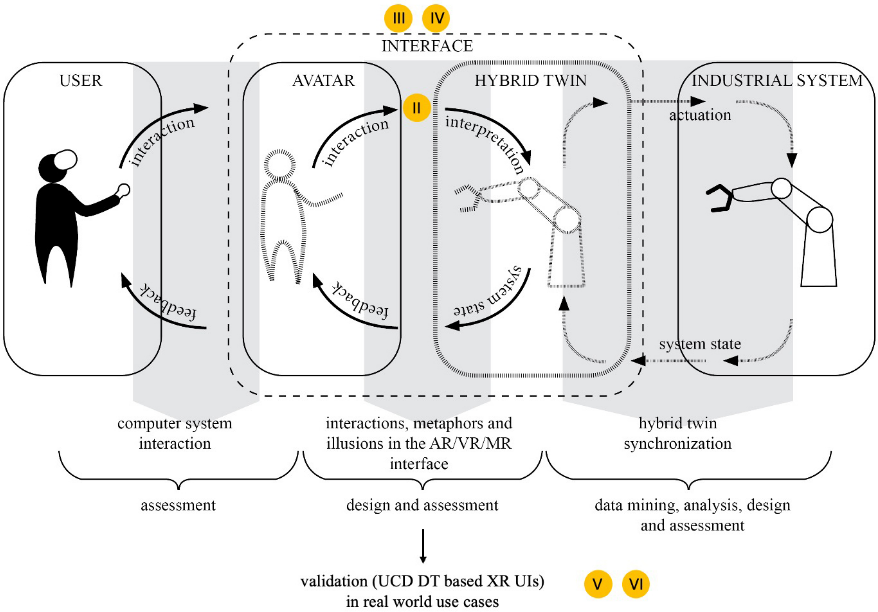

Figure 8: DT and HCI loops integration diagram

3.1 Programmatic design, assessment and validation based on I paper

The previous image shows how the testing and validation presented in II, III and IV are allocated in the conceptual diagram. To give a brief summary:

- The pilot study presented in II tackles, at least partially, the type of interactions happening within the digital environment by comparing three type of interactions. The metrics and evaluation tools used are NasaTLX, SUS and UEQ, speak aloud protocol and time to complete the task. A total of only 6 subject participated in the study due to COVID-19 restrictions (15 planned).

- A comparison of DT based VR interface and native robot teach pendant on a pick and place task is presented in III. The metrics and evaluation tools used are NasaTLX, Goodspeed questionnaire, eye tracking in VR, heart rate, and time to complete the task. A total of 40 subjects took part in the experiment.

- A comparative evaluation of DT based XR UIs (AR and VR) using hand manipulation and robot teach pendant is presented in IV. The metrics and evaluation tools used are NasaTLX, SUS and UEQ, speak aloud protocol and time to complete the task. In a first session 21 subjects took part in the experiments. A second session tested the system on another group of 21 subjects. During the second session path point accuracy has been collected. The second session has not been validated by publication.

The works presented in V and VI attempt a pilot integration and validation of the proposed user centric design and evaluation approach using XR UIs in real world applications. A total of 6 subjects took part in the evaluation of the application presented in VI.

3.1.1 Subjective metrics

Before presenting the above mentioned studies it is worth introducing the adopted subjective questionnaire metrics at once. These have been selected among the available qualitative evaluation tools based on in literature and previous studies in the HRC field.

System Usability Scale

The system usability scale is a popular 10 items subjective assessment questionnaire designed in the eighties and presented in [23]. The questionnaire aims at detecting higher level usability of the system, which makes it adaptable and usable to compare different types of systems but at the same time quite generic. The proposed statements are rated on a five point Likert scale with a positive and negative end. The usability ranking happens on a scale from 0 to 100 and assessing the system usability as shown in Figure 9.

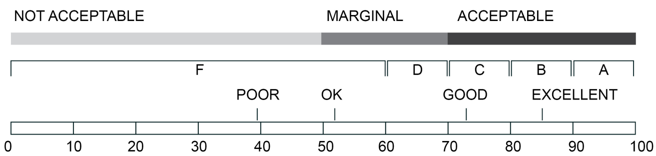

Figure 9: System Usability Scale Ranking

Nasa Task Load Index

NasaTLX [49] has a long adoption history both in research and other domains, not exclusively industry. The task load evaluation questionnaire aims at detecting the perceived workload and effort while performing a task along seven dimensions, namely: mental demand, physical demand, temporal demand, performance, effort and frustration. These are ranked on a scale form 0 to 100, very low to very high, with intervals of 5. In the following studies we use the raw scale results without adopting the sub-scales and ranking.

User Experience Questionnaire

The User Experience Questionnaire is a much newer tool [78] allowing the assessment of subjective experience and usability aspects while performing a task by means of an interactive product. The long version has 26 pairs of contrasting attributes ranked on a 7 point Likert scale. This questionnaire aims at detecting different dimensions of the interaction experience including attractiveness, perspicuity, efficiency, dependability, stimulation and novelty. For the following studies we decided to use the short version which is focusing on hedonic and pragmatic qualities with 8 pairs or attributes only. There are two main motivations for this choice. The first is that the questionnaire is used together with other assessment methods and we want to avoid overloading the users with the evaluation tasks. The second is that most of the experimental protocols include repetitive measures for different types of interactions or hardware devices which would result in longer and tiring sessions. Regardless, the UEQ scores results on a scale between 3 and -3 with values between 1 and -1 being average and above 2 or below -2 resulting in a good or bad user experience during the interaction.

Godspeed Questionnaire

The godspeed questionnaire [17] is a validated subjective questionnaire aimed at assessing user perception of robotic systems, industrial or service robots, on different concepts namely anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety. These are scored by a set of pairs attributes on a 5 point scale. The final scoring for each concept is taken by averaging results of the items.

4 Assessment and evaluation of DT based XR interfaces for HRI

4.1 Interaction methods in XR, a pilot study (II)

We aim at an evaluation of user experience and performance of different types of interaction paradigms within the DT based XR interface by means of comparing three types of controller based interaction methods in a virtual robot pick and place task.

This topic has been and still is largely explored in many studies [68]. A few studies focus on the evaluation of virtual UIs data selection [105], data entry [143], and visualisation [43]. Other examples try to assess the performance of different interaction devices for the specific task [69] or the impact of user perception while using different type of controllers [12]. As already mentioned the type of interaction paradigms and metaphors possible within the simulated environment is practically unlimited, yet restricted by the type of hardware and technologies mediating them without the virtual scenario.

4.1.1 Hardware and software architecture

The virtual scenario adopted in the study and relative software architecture for DT twinning is based on the work presented in [75]. The digital replica of the robot setup and laboratory is developed by using Unity game engine [8] and custom scripts to allow the real-digital equipment synchronization. The focus of this study is on the Yaskawa Motoman GP8 robot model and virtual interface, Figure 10.

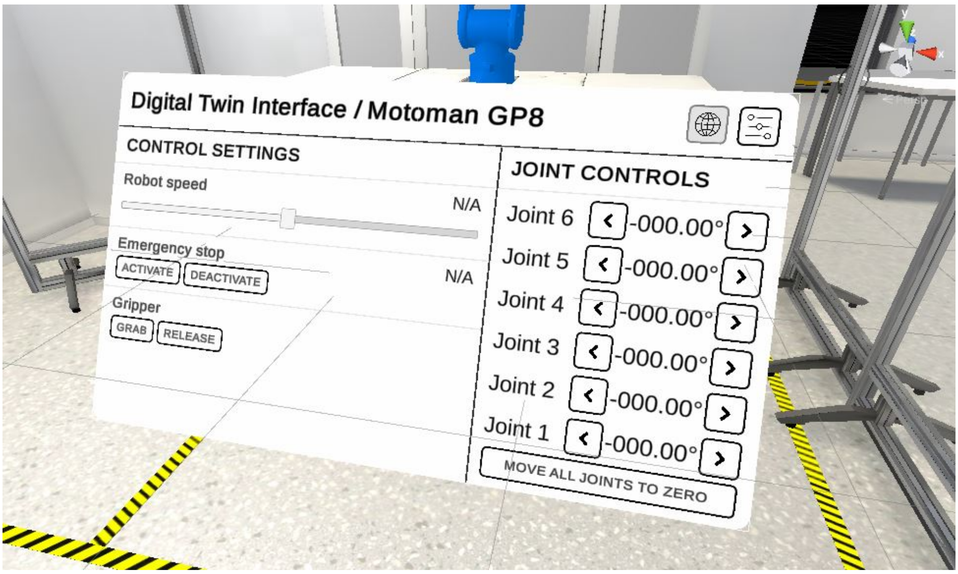

Figure 10: DT virtual user interface

The interface comprises different functions and control settings allowing the operator to directly manipulate each joint rotation and speed, activate or deactivate the gripper, move the robot to home position and switch from synchronised to virtual mode. Each joint rotation angle is shown in the interface and dynamically changed during the movement. The study is focusing on the interaction methods within the virtual scenario therefore we don't aim at controlling the real robot at this point but its virtual replica only.

Oculus Rift S and controllers, shown in Figure 11, is used for the visualization and interaction with the virtual environment.

Figure 11: Oculus Rift S headset and controllers

This specific headset is not standalone but connected to a local PC workstation equipped with Intel core I7-6700HQ, Nvidia GEFORCE GTX 970m graphic processing unit and 16 GB of ram.

4.1.2 Interaction methods

Three types of interaction methods are compared during the assessment. The first adopts a controller based navigation and interaction paradigm (CBI). This leverages on the Unity input system software package (v1.0.2) granting hardware agnostic action maps creation that can be bound to one or more input values. This method allows the system to adopt different hardware input (game controllers, keyboard, XR controllers) to the same virtual UI actions. Visually this works as old style desktop based UI interfaces navigation and selection. The user can browse through the interactable interface elements, highlighted upon selection, by means of the controller joystick. The element can be clicked by pressing the controller trigger button. The joystick and trigger button are shown in Figure 12.

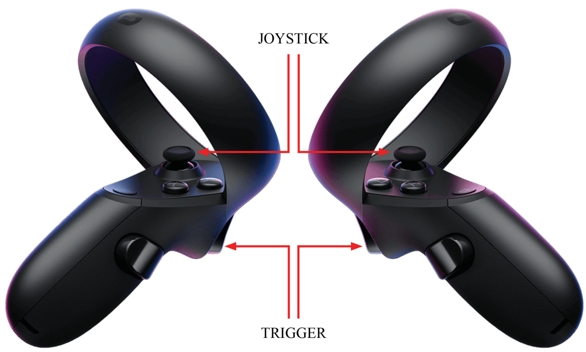

Figure 12: Oculus Rift S controller trigger button and joystick

Table 3 shows the hardware input mapping for each type of UI interaction in the virtual scenario.

The second method is also well known type of interaction paradigm in XR applications, namely a ray cast pointer (RCI). The ray cast pointer is triggered when the controllers are directed towards any interface in the scene. Consequently the UI components (buttons, sliders etc.) are highlighted when hit by the ray. The selection happens by pressing the

Table 3: *Input mapping for XR controllers*

Action	Controller binding
UI Navigation	Right or Left XR controller Joystick
Click	Right or Left trigger button pressed
Slider	Right or Left XR controller Joystick (left and right only)

controller trigger button. Both Oculus Rift controllers can be used interchangeably during the task. Figure 13 shows the RCI while interacting with the robot speed slider.

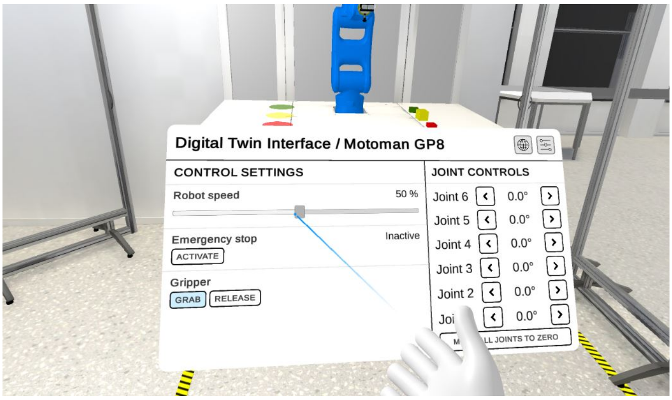

Figure 13: Ray cast pointer interaction method

The third interaction method adopted in the study is a direct interaction (DI) with the virtual UI. This is mediated by an avatar hand (non dynamic) which can be used to press the interface buttons directly as if they were in a physical touch screen interface. The method adopts an invisible ray cast pointer that triggers interaction with the interactable elements only when at a specific distance from the UI (less than 5mm). The ray is also always directed perpendicularly to the interface itself to optimise the interaction. Custom scripts are allowing the interaction logic integration with the Unity environment. Figure 14 shows the DI while interacting with the robot UI.

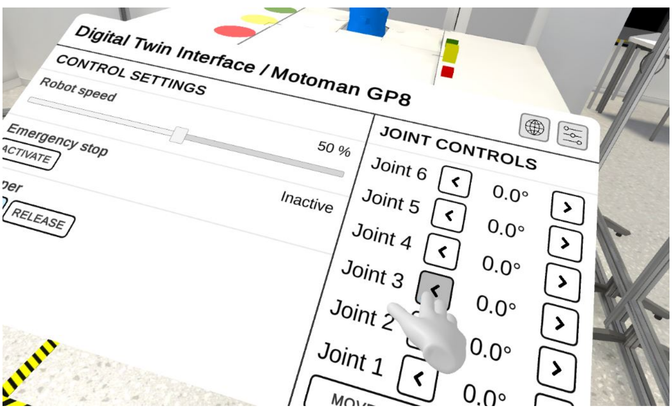

Figure 14: Direct interaction method

4.1.3 Study protocol

As already mentioned the study aimed at involving 15 participants in the assessment but due to COVID-19 restrictions only 6 people were recruited. Each subject performed the task three times by using the mentioned interaction methods (CBI, RCI, DI). The sessions are counterbalanced so that each subject is involved in a different sequence of interactions. The tasks consist in picking three cubes (red, green and yellow) and placing them in the correct positions highlighted on the robot workstation. This task involves different and repeated interactions with the UI components mainly the joint controls and gripper area. Figure 15 shows the experimental protocol sequence for each subject.

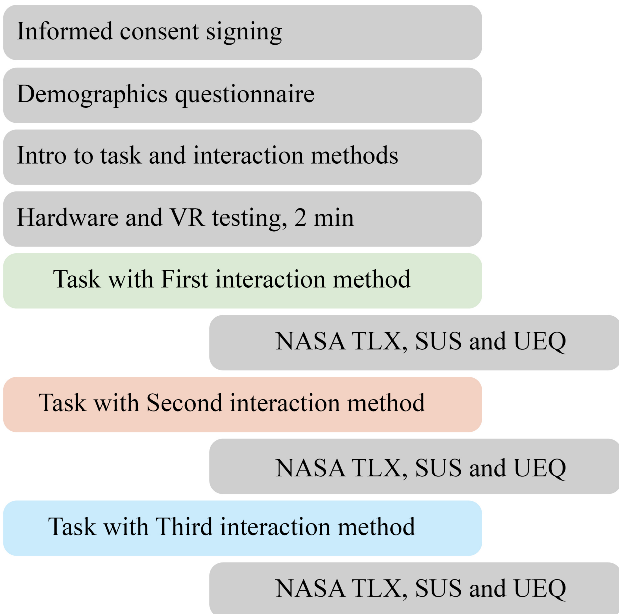

Figure 15: Experimental protocol

Prior to each session, participants are presented with an overview of the main objectives of the study, the required tasks as well as instructions on navigating the virtual reality user interface and the types of interactions tested during the experiment. A demographics questionnaire is also administered to collect basic user data and assess the level of acquaintance with XR equipment, related applications and industrial robot control. This questionnaire is based on a 5 point Likert scale having a positive and negative extremes. Regardless of their prior experience with VR navigation and equipment, each participant is allowed to familiarize with the head-mounted display and controllers for two minutes. Following each session, participants completed three questionnaires assessing workload, system usability, and user experience for each interaction method. For this purpose we adopted SUS, NasaTLX and UEQ questionnaires. The subjective metrics paired with performance in completing the task, measured by the total time spent in replacing the cubes. A speak aloud protocol was also included and comments from the participants noted down during each session.

4.1.4 Results

The experimental group consists of 4 male and 2 female participants with an average age of 30 (SD ± 9) years old. The education level varies from PhD (3 subjects), to master diploma (1 subject) and bachelor diploma (2 subjects). Acquaintance with HMDs VR technology is generally high except for one subject having no experience with neither equipment nor applications related to XR. Nevertheless, all subjects have previous experiences with gaming consoles and controllers. Two participants have extensive knowledge of industrial robot control and programming. Results for the other 4 subjects are evenly distributed along the measuring scale with only 1 participants having absolutely no prior experience with industrial robot control and programming.

Performance

Measurement of performance demonstrate that the most efficient XR interaction for the given task is the RCI method with 195.5 seconds (SD ± 127.19). Following is the CBI with 312.33 seconds (SD ± 115.5) and DI with a higher value of 351.17 seconds (SD ± 138.16). Table 4 provides an overview of the results.

Table 4: Task performance

Method	Mean (s)	SD (s)
RCI	195.5	± 127.19
CBI	312.33	± 115.5
DI	351.17	± 138.16

Workload, usability and user experience

Results for the NasaTLX questionnaire show that the perceived effort for each session is in general quite low. RCI shows the best overall score, 18.61, followed by CBI, 26.95 and DI, 36.81. CBI has above average scores for Mental and and Temporal demand. Higher scores in these items are also evident in the DI method results. Table 5 shows an overview of NasaTLX raw scores for each questionnaire metric.

Table 5: NasaTLX raw scores results

	RCI	CBI	DI
Mental	25.0	42.50	42.50
Physical	12.50	17.50	33.33
Temporal	18.33	27.50	44.17
Performance	24.17	30.83	33.33
Effort	20.00	18.33	34.17
Frustration	11.67	20.83	33.33
Overall	18.61	26.25	36.81

System Usability Scale questionnaire, Table 6, show optimal usability for the RCI with a score of 90.83 (SD ± 7.31). These values are followed by both CBI and DI methods with respective results of 72.91 (SD ± 15.97) and 70.83 (SD ± 12.72) positioning the two interaction methods on a good level of usability.

Table 6: System Usability Scale questionnaire results

Method	Mean (s)	SD
RCI	90.83	$\pm$7.31
CBI	72.91	$\pm$15.96
DI	70.83	$\pm$12.72

User experience is perceived as really good, for both pragmatic and hedonic aspects, for the RCI with an overall score of 2.06. Positive scores are reported for both CBI and DI methods with values of 1.15 and 1.25 respectively. As expected the CBI has the lowest score for hedonic qualities. Table 7 shows the UEQ results overview.

Table 7: User Experience Questionnaire results

	Pragmatic	Hedonic	Overall
RCI	2.46	1.66	2.06
CBI	1.67	0.65	1.15
DI	1.25	1.25	1.25

Speak aloud protocol

Users provided a few comments on the different type of interaction methods. In general the CBI seemed a bit outdated and not novel, but quite easy to manage, because of familiarity with similar systems. One user, in particular, felt that this method was showing a higher responsiveness of robot joint controls. DI was perceived as physical demanding in two cases, because it required to get closer to the interface and interact directly with the controls. A few users also found the direct interaction being too sensitive for the task. RCI was rated as the most efficient and less demanding interaction paradigm overall. Other general comments on the system highlighted the necessity of having adequate feedback on the robot joint position and grabbing action. A few remarks related to the comfort of wearing the headset, and position of specific buttons on the main virtual interface.

4.1.5 Discussion

Preliminary test show that the higher perceived usability and positive experience with traditional ray-cast based interaction. This is also confirmed by lower level of effort and higher performance figures. Lower effort might be related to the fact that the interaction is happening at distance without need of physical movements. This is also confirmed by the low task load results for the controller based interaction. A larger subject sample could confirm if higher scores for the RCI are due to previous experience with similar VR systems and controllers. The adoption of modern HMD with hand tracking capabilities might also be investigated. Further experiments described in the upcoming sections try to compare DT based XR UIs with the specific robot controls in terms of usability, efficiency, task load and user experience.

4.2 Development and evaluation of DT based VR UI with teach pendant on pick and place task (III)

The third study aims at comparing industrial robot control task through a DT based VR interface and native teach pendant. The objective is to understand if there is any performance and user experience degradation while performing the task in XR or if these can be a reliable substitute for native machine controls.

4.2.1 System architecture and components

The experiments make use of the same DT system adopted in II presented in [75] and shown in Figure 16. The robot brand and model is Yaskawa Motoman GP8. The DT system provides the same joint rotation accuracy as the real one. As already mentioned the virtual user interface allows to control the robot in coupled or virtual mode. In coupled mode both virtual and real robots are controlled simultaneously though local network connection. In virtual mode the real robot is disconnected and only the digital replica manipulated in the virtual scenario. Moreover, the interface allows to either directly control the robot joints rotation and position, or create a planning task to be performed in a second stage. For this study a virtual mode and direct control methods are chosen. Furthermore a custom script is introducing network instability and lag in a random sort into the XR interface scenario.

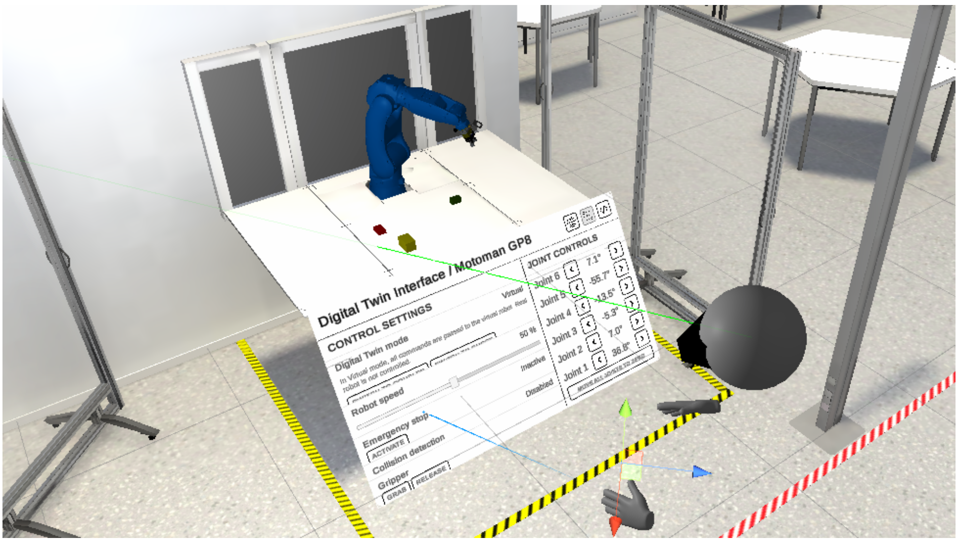

Figure 16: DT base UI setup

A custom attention tracker system is developed for the virtual environment with the aim of recording the behaviour of the subjects while performing the task. A Vive Pro Eye headset is used for this experimental session. The native Vive SRanipal SDK is employed to access eye tracking data in Unity which is typically provided in quaternions. The approximation of the user eye gaze is utilized to detect interactions with the objects of interest and register attention events. This method isn't confined to a particular HMD model and can be replicated across various eye tracking systems. The attention tracking module consists of different software components shown in Figure 17, including:

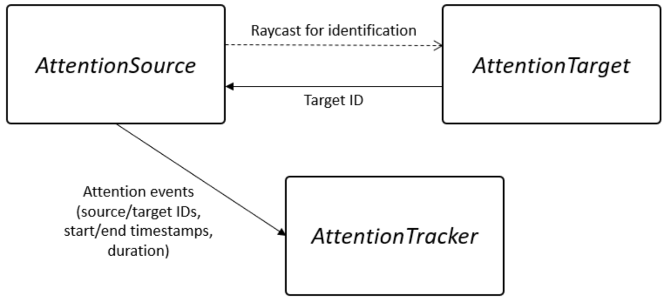

Figure 17: Attention tracker component architecture approximating eye tracking within the XR interface

- The Attention Target script is responsible to identify any object or interface component the script is assigned to as object of interest for the Attention Source.
- The Attention Source is responsible to detect interactions with the Attention Targets and save duration and timestamps (milliseconds) for the beginning and end of these interactions. The minimum precision is 8.3 ms as the environment is executed at 120Hz. The detection of interaction is based on a raycasting from the normal surface of the HMD. When an Attention Target is in the line of sight of this invisible ray, a new attention event is triggered and data saved into a separate session files.
- The Attention Tracker script is responsible for the management of attention events (start and stop the sessions) data storage and export in JSON format.

The main target of the attention tracking system is the virtual UI as we aim at estimating the effectiveness of design and ease of use by averaging the amount of time spent in each section. These information can be used in a later stage to improve controls positioning and cues within the virtual UI. For this reason the interface is divided into three areas corresponding to the joint control section, the general control section and the interface header, Figure 18. Attention Targets are assigned to each of the interface sub sections. Attention Targets are also allocated for each virtual robot joint.

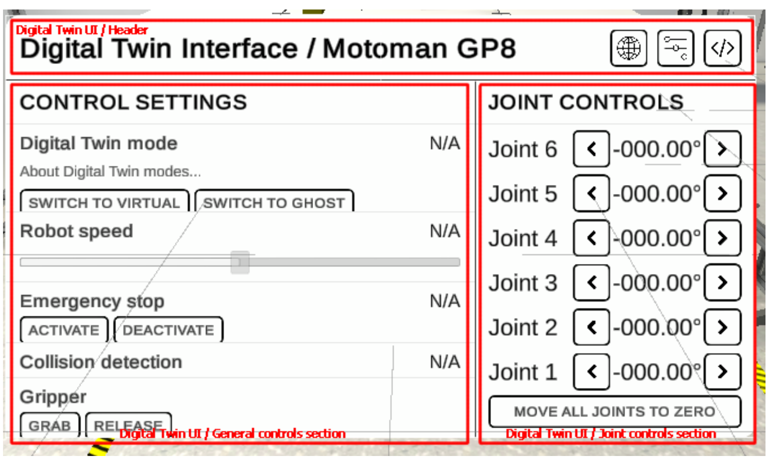

Figure 18: Attention Target sections in the VR UI

As already mentioned the hardware components consist of Yaskawa Motoman GP8 robotic arm and native teach pendant. The robot is mounted on a work bench which is also replicated in the virtual environment. The XR visualisation and interaction hardware used is a Vive Pro Eye with relative controllers and tracking sensors. Vive Pro is not a standalone HMD and is therefore connected to a local workstation on which the DT application runs.

4.2.2 Experimental protocol and metrics

The experimental protocol includes of a simple pick and place task defined as follows. The target objects are three wooden cubes of different sizes, that the operator is required to move from an initial location to a target area. The pick-up locations are different for each cube and identical in the real and virtual scenario. The target location is highlighted both on the real and virtual workbench. Every participant have to interact with the robot teach pendant and VR UI robot joints controls to reach the cubes, and with the gripper actuator to be able to grab them. This last task requires higher precision and an initial trial and error phase, especially in real world scenario, which is expected to improve during the execution. A total of 40 participants performed the task divided in two groups to counterbalance the experimental sequence of interface interactions. A short amount of time was given to each user to get familiar with the UIs while a full presentation of the research objectives, materials and methods, risk, benefits, data handling and privacy related to the study was presented. Each user signed an informed consent before starting the session.

Prior to the task generic demographics of the operators, including age, gender and nationality, are collected though a custom questionnaire. As previous experience with robotics, including level of exposure and acquaintance with related tasks, might have influence on expectations towards the specific system, another questionnaire detects these subjective metrics. A combination of qualitative and quantitative metrics are employed to assess the human robot interaction while performing the task with the two different interfaces. The quantifiable objective measures are aimed at capturing the user task performance, while at the same time collecting information on the focus of attention within the DT based VR UI. The total time spend in completing the task is collected both in the virtual and physical setup. Within the immersive environment the following additional data is collected:

- Average focus of attention duration with each virtual interface section, as explained in the previous paragraph.
- Average focus of attention duration focusing on each of the robot joints.
- Average focus of attention doing something else.

Two post task subjective surveys are employed to assess workload while performing the task for each type of interface (NasaTLX), and assess the users perception of robots (Godspeed Questionnaire).

4.2.3 Results

Demographics

The first subject group (A) consisted of a total of 20 participants, 16 male and 4 female, with an average age of 29.5 years old, and backgrounds varying from engineering, business administration, bachelors and master students, researchers and lecturers in engineering disciplines. The requested self assessment on robotics systems knowledge and skills was scored on a ten point scale, with 1 corresponding to no experience and 10 to expert in robotics, with an average result of 3.9 therefore showing low level of expertise in this field. The second subject group (B) consisted of 17 male and 3 female participants with and average age of 29.9 years old and backgrounds including students, researchers and professors in the mechanical engineering domain. Likewise the average results for the robotics skills self assessment gave a low score with an average of 4.2.

Performance

Results related to the total average time in placing each cube in the target area show a trend of longer timing for the real world setup compared to the DT based UI. This is true for both subject groups. Nevertheless, a significant higher performance time with the robot native teach pendant in noticeable in group B. This might imply that XR environment might work as a training for real world setup or that the level of acquaintance for the robotic systems was underestimated. Table 8 and 9 show the results for the two groups.

Table 8: Average task completion time with the physical and virtual interface for each cube in seconds, group A

	Physical	**Virtual**
Cube 1	225	144
Cube 2	210	115
Cube 3	184	105

Table 9: Average task completion time with the virtual and physical interface for each cube in seconds, group B

	Virtual	**Physical**
Cube 1	114	178
Cube 2	92	130
Cube 3	61	159

A further assessment of this trend happened by testing a subset of 4 subjects with different expertise in robotics in two consecutive trials using the VR interface. Results, Table 10 show the lack of correlation between expertise in robotics and task performance in the XR scenario.

This performance discrepancy might be due to other factors like expertise with XR systems or other type of robotics equipment. This should be investigated further in other experimental sessions. Statistical analysis confirms the absence of correlation between time to perform the task in VR and expertise with robots while, as expected, there is a strong negative correlation between performance time while interacting with the tech pendant and expertise with robotic systems, r_s=-0.49 $p<0.01$. In general the time to complete the replacement of each cube is consistently decreasing for both the VR and physical interaction setup as shown in Figure 19.

Questionnaires responses

Table 11 reports the results for the Godspeed Questionnaire assessment of perception of robotic system for group A and B. As already mentioned each item is scored on a 5 point Likert scale between two different definitions.

Results for this first questionnaire show similar trends for both groups. There are

Table 10: *Task completion time (in seconds) for the VR interface of a subset of users with different expertise in robotics*

Trial 1				
Expertise with robotics	2	9	1	2
Cube 1	360	120	180	120
Cube 2	240	120	120	90
Cube 3	60	180	120	90
Trial 2				
Cube 1	300	120	60	60
Cube 2	240	60	60	45
Cube 3	60	45	60	45

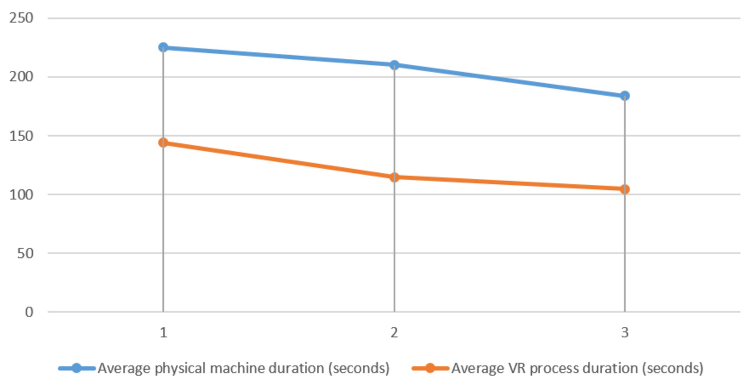

Figure 19: Comparison of replacement time per cube with VR and teach pendant interfaces.

no significant noticeable differences between the users perception of the robotic system while performing the task with the physical and virtual user interfaces. The virtual setup is considered more interactive and responsive than the real one with a trend of general higher perceived intelligence in all sub-scales items. This is confirmed by total average values for VR, 3.37, and physical robot, 3.01. Moreover, the total averages for perception of safety of the VR interfaces is higher (3.6) than the relative value for the physical setup (3.2).

Results for the NasaTLX questionnaire are shown in Table 12. Results for the group A show higher values of mental demand, perceived effort and frustration for the VR interface. This is in contrast with results of the second group having reversed trends for all the sub-scales. These results can be correlated to the order of interface interaction. The first interface every user interacted with is perceived as less demanding, while giving lower performance results, especially in group A. The overall perceived task load is higher for the physical machine setup in the second group. There are no significant statistical correlations between NasaTLX results and heart rate results reported below. There is a positive correlation between task load and total average time spent in completing the task for both the VR UI, $r=0.47$ $p<0.01$, and the physical setup $r=0.31$ $p<0.05$. No correlation is noticeable between task load and expertise with robotics and Godspeed questionnaire.

Table 11: Godspeed questionnaire results for VR and physical robot interface setup, for group A and group B

	Group A		Group B	
	Physical	Virtual	Virtual	Physical
Anthropomorphism (mean)	2.97	2.47	2.67	2.8
Fake-Natural	4	2.65	3.55	3.45
Machine-like-Human-like	2.6	2.5	1.95	1.9
Unconscious-Conscious	2.45	2.3	2.35	2.3
Artificial-Lifelike	2.55	2.15	2.45	2.2
Moving Rigidly-Moving Elegantly	3.25	2.75	3.05	3.45
Animacy (mean)	2.68	2.7	3.2	2.9
Dead-Alive	2.5	2.4	2.9	2.45
Stagnant-Lively	2.85	2.85	3.2	3
Mechanical-Organic	2.3	2.35	2.25	1.85
Artificial-Lifelike	2.3	1.9	2.4	2.05
Inert-Interactive	2.75	3.35	4.1	3.35
Apathetic-Responsive	3.4	3.45	4.25	3.7
Likeability (mean)	3.69	3.46	4.07	3.6
Dislike-Like	4	3.5	4.45	3.8
Unfriendly-Friendly	3.55	3.3	3.95	3.1
Unkind-Kind	3.4	3.35	3.7	3.1
Unpleasant-Pleasant	3.6	3.55	3.95	3.35
Awful-Nice	3.9	3.55	4.3	3.65
Perceived Intelligence (mean)	3.12	3.34	3.39	3.2
Incompetent-Competent	3.2	3.3	3.65	3.2
Ignorant-Knowledgeable	2.95	3.45	3.54	3.05
Irresponsible-Responsible	3.45	3.5	3.3	3.15
Unintelligent-Intelligent	2.9	3.2	3.2	2.8
Foolish-Sensible	3.1	3.25	3.35	3.15
Perceived Safety (mean)	3.6	3.5	3.7	3.0
Anxious-Relaxed	4.1	3.55	3.7	2.6
Agitated-Calm	3.85	3.9	3.8	2.95
Quiescent-Surprised	2.95	3.05	3.7	3

Table 12: NasaTLX results for subject group A and B

	Group A		Group B	
	Physical	Virtual	Virtual	Physical
Mental Demand	45.5	55	43.5	60
Physical Demand	29.2	25.7	22.5	46.5
Temporal Demand	43.5	37.5	40.7	55.5
Performance	31.25	35	38.5	45.5
Effort	34.25	43.25	40.7	62.5
Frustration	31.75	41.25	21	47.5
OVERALL	35,91	39,62	34,48	52,92

Attention tracker results

Results for the eye tracking data collected during the interactions with the VR setup show an average higher time spend looking at the virtual UI than the robot joints. Figure 20 and 21 show the results for group A and B in form of histograms.

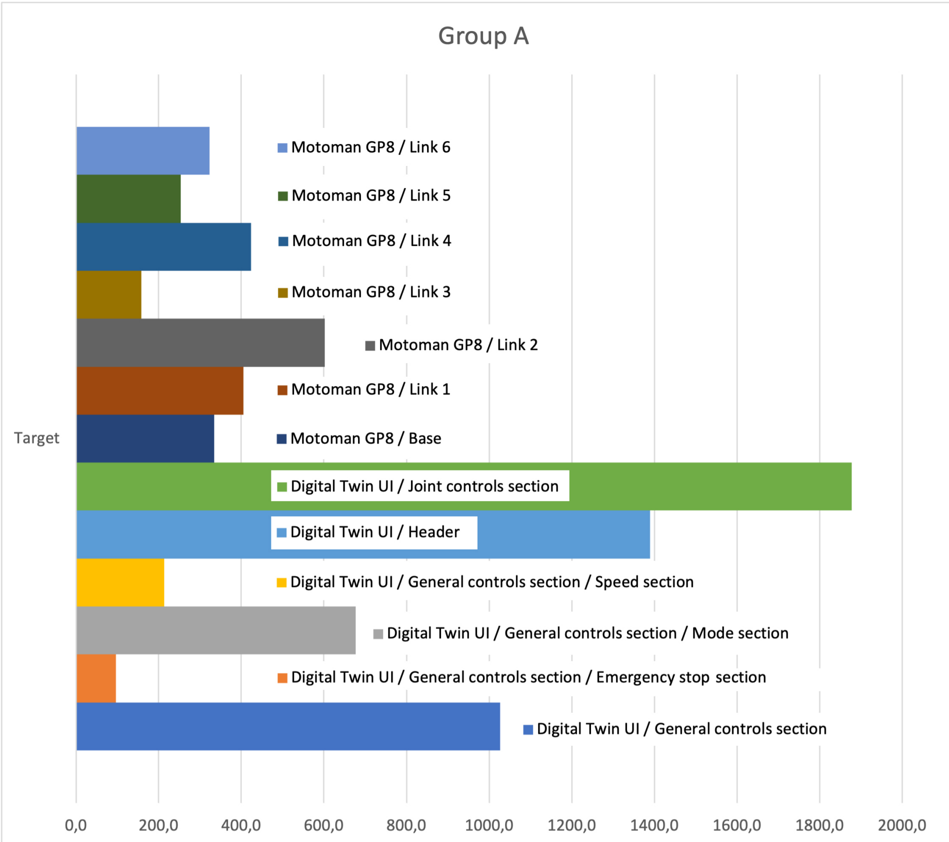

Figure 20: Attention tracking results in seconds per target, group A

The two subgroups slightly differ in the specific regions of interests. Group A spent more time on general controls, joint controls and header area. Group B spend similarly an high amount of time in joint control section, but the attention on the header and the general controls was definitely lower. Higher time is spent by participants in group B on the robot Joint 6. Tables with timing (in seconds) spent looking at DT UI sections, robot joints and operators avatar hand for each participants are reported in the III annex. The higher time results for the robot Joint 6 is quite understandable as it represents the gripper joint involved in the pick and place task. No other clear correlations can be made on the level of expertise with robotic systems.

To have a better overview of the results and capture the general trends of the attention tracking system the data is normalized. Plots 22 and 23 show a cleared overview of the focus of attention of the operators. The two plots report the time spent looking at the UI elements and robot joints. The labels have to be interpreted as follow: GS, general control section; ES, emergency stop section; MS, mode selection; SS, speed selection; JC, joint controls; H, header; L 1 to 6 represent the robot links/joints. The plots report the

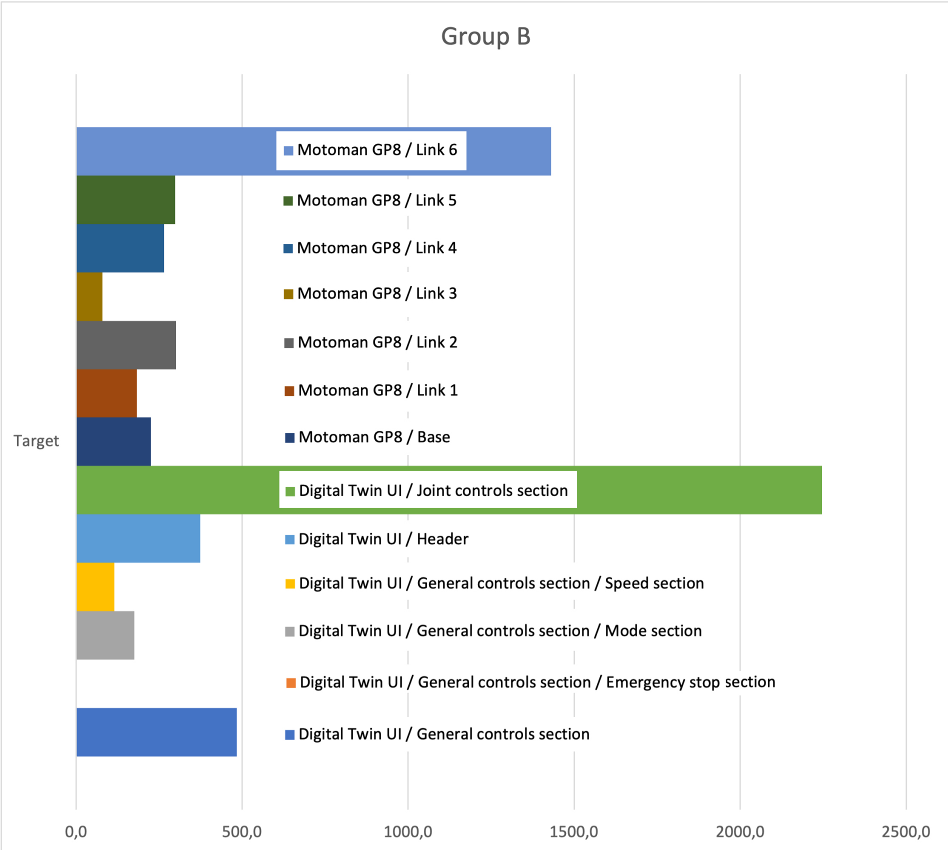

Figure 21: Attention tracking results in seconds per target, group B

mean, median, maximum and minimum value for each VR UI section. The two images show a higher amount of time spent in looking at the interface header, followed by the joint control section. The first result does not have explanations related to any necessary actions to be undertaken in the VR environment while can be explained by an inaccurate estimation of the eyesight. Improvements have to be made on the focus of attention tracking system to avoid detecting false UI target elements such as the header, while, for instance, the operator was actually interacting with the joint controls. Accuracy can be also improved by correlating eyesight with the position and interactions of the ray cast pointer used to control the UI elements and detecting the time spend for each interaction.

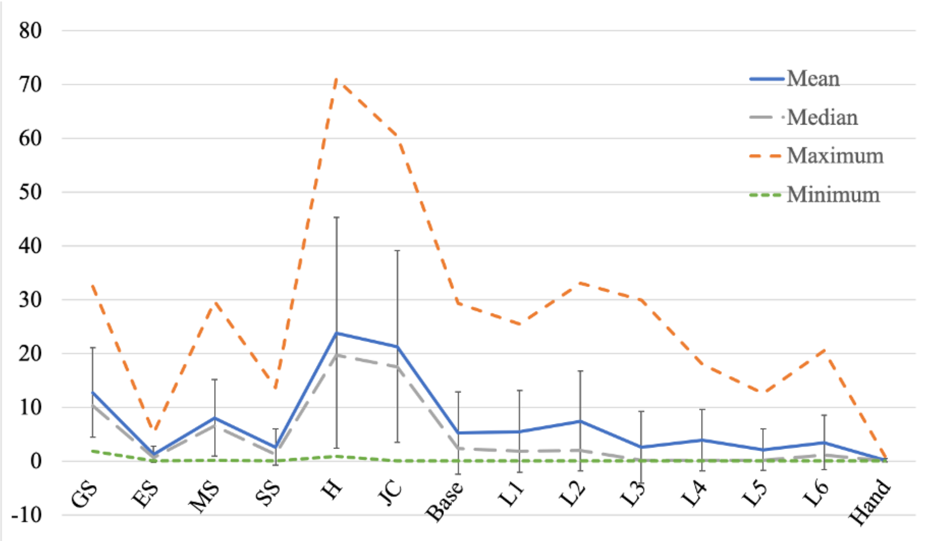

Figure 22: Attention tracking normalised results per target, group A. The error bars represent a single standard deviation from the mean.

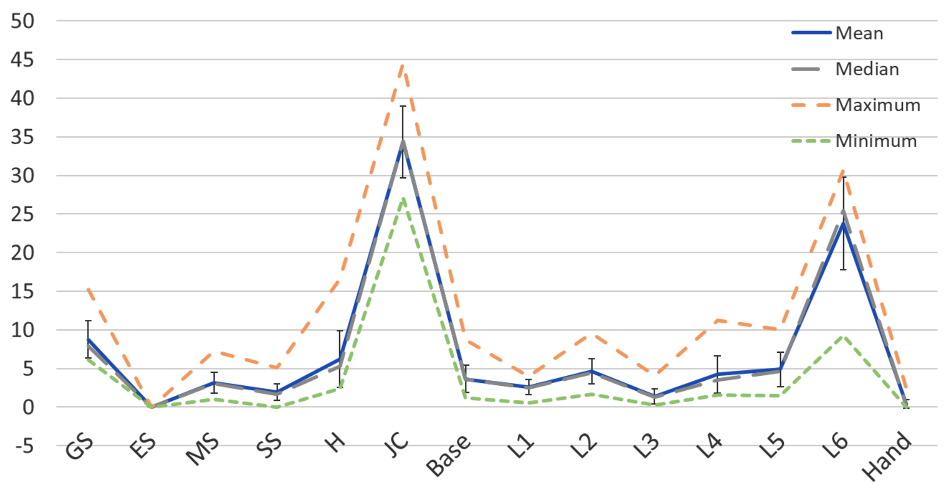

Figure 23: Attention tracking normalised results per target, group B. The error bars represent a single standard deviation from the mean.

Physiological stress monitoring

Heart rate for every participant was collected during each session by means of a commercial grade writs band. Results for this metric are inconclusive and don't show any significant correlation with the other variables. Average heart rate is almost identical for the VR UI setup and physical interface in group B, while slightly higher in the latter case for participants group A. Table 13 shows the results for each subgroup and type of user interface interaction.

Table 13: Maximum, minimum, and average heart rate values for each type of interface interaction in group A and B

	Max (BPM)	**Min (BPM)**	**Average (BPM)**	**SD**
		Group A		
Physical	117	75	90.05	$\pm$11.99
Virtual	105	76	85.05	$\pm$7.53
		Group B		
Virtual	99	83	90.55	$\pm$4.27
Physical	100	79	90.1	$\pm$6.66

4.2.4 Discussion

Overall results show that the DT based UI is largely comparable to controlling the robot with the native teach pendant with a considerable higher performance of the VR setup, which in turns is not always the least demanding method, especially in terms of frustration and mental demand. The system shows several flaws regarding the eye tracking system setup and physiological response which need to be addressed in future studies.

4.3 Development and comparative evaluation of DT based XR UI with teach pendant on a path planning task (IV)

The work presents an evaluation of DT based XR interfaces for path planning task. The study aims at assessing usability of AR and VR interfaces in comparison with native teach pendants for the control of two different industrial robots. The publication reports preliminary results for the first sessions of experiments and a not yet published and validated report of the second session.

4.3.1 Hardware and software architecture

The hardware setup makes use of two different industrial robots, ABB IRB 1600 and ABB IRB 1200 located in two different locations namely Tallinn University of Technology (Estonia) and Technical University of the Shannon (Ireland).

The interaction and visualization hardware setup for the XR UIs consist of two headsets for virtual and augmented reality, Meta Quest 2 and Microsoft HoloLens. Both support inside-out camera based positioning and hand tracking, with no need for external sensors, and both are standalone.

The software architecture integrates ROS Industrial [5] standard with MoveIt [29] motion planning framework. The first allows the system to be expanded to other robot models supported by ROS-I while maintaining the same UI. The architecture shown in Figure 24 consists of two main sides connected thought local network.

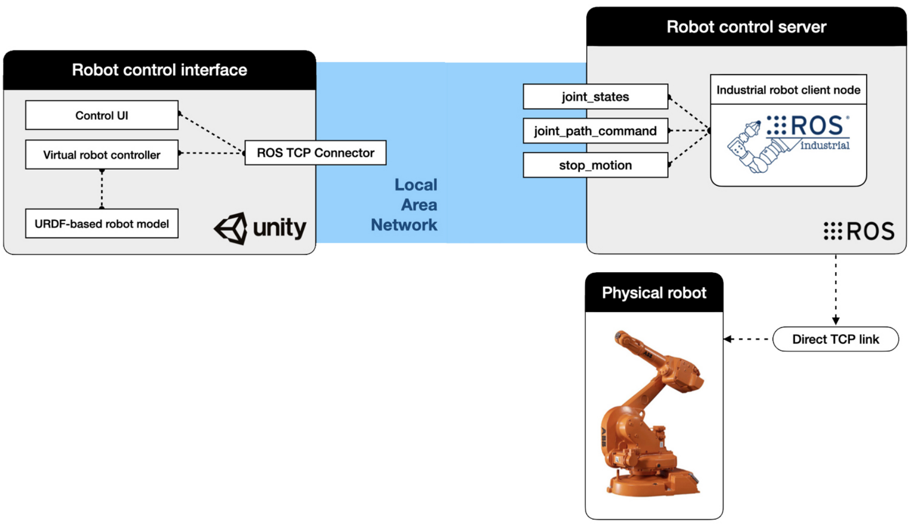

Figure 24: Unity ROS Architecture

- The Unity application side takes care of the immersive environment visualization, integration of interactions modality and interface components to control the robots. These are programmed in Unity though native SDKs and custom C# scripts. Communication with ROS happens by adopting components from the Unity Robotics Hub projects and customs scripts for handling MoveIt commands independently from the actual robots connected to ROS. The robot models are imported by means of Universal Robot Description format and ROS package compatible with the ROS-I standard.

- The ROS side takes care of the control of the real robot joints positions, while up to date robot poses, joint state, and commands are communicated to the Unity application though local network.

ROS runs on a Linux based dedicated server, while the Unity application is installed directly on the headset devices. The positioning of the AR content in real world coordinate system is possible by adopting Vuforia SDK for Unity and based on image targeting thought the device cameras.

4.3.2 XR UI and interaction methods

The type of interactions and user interfaces adopted for the systems rely on native Unity UI packages and two different SDKs namely, Microsoft Mixed Reality Toolkit (MRTK) [95] and Oculus Integration. These allow the adoption of hand avatars into the augmented or virtual reality scenarios and natural poke or grab interactions with either UI elements (buttons, check boxes or drop down menus) and the path planning tool. Figure 25 and Figure 26 show the virtual user interface functions and robot setup for the experiment with the path planning tool.

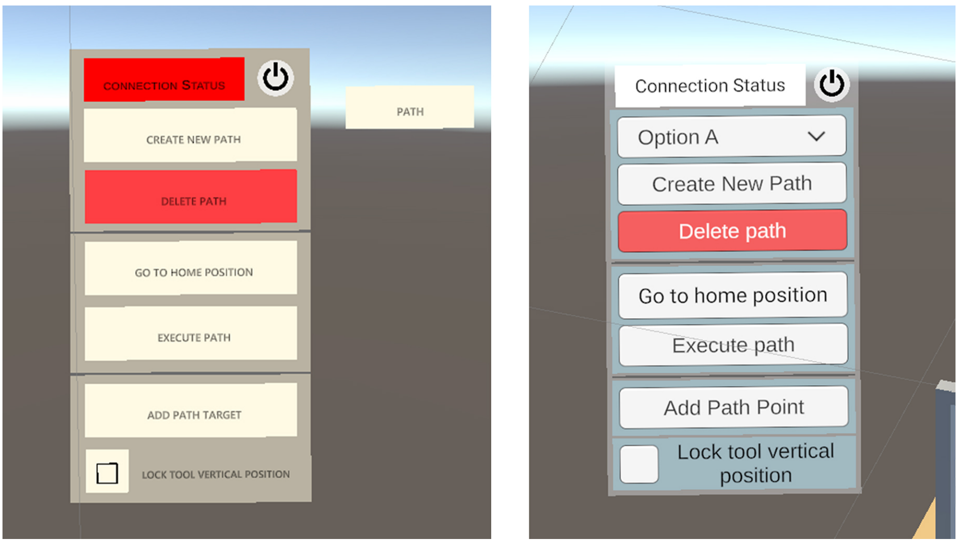

Figure 25: XR UI in VR and AR applications

The user interface allows creating new routines and path points for each required task. Paths, path points position and rotation are saved on the headsets memories in JSON format for further evaluation and comparison with real world path points. The interface allows setting the robot position to home and executing the created routines. When connected to the ROS server both real robot and DT move in a synchronized way.

To create the points the user can manipulated the sphere (path planning tool), shown in Figure 26, and press the adjacent plus button. Path points are spawned at the center of the sphere and visualized as green dots connected by segment lines to each other. Coordinate system axes are provided at the center of the path planning tool to facilitate the positioning and creation of the points.

The path and metal stand in front of the robot are modelled to be identical to the real world scenario. The real world user interface is the ABB DSQC679 teach pendant for

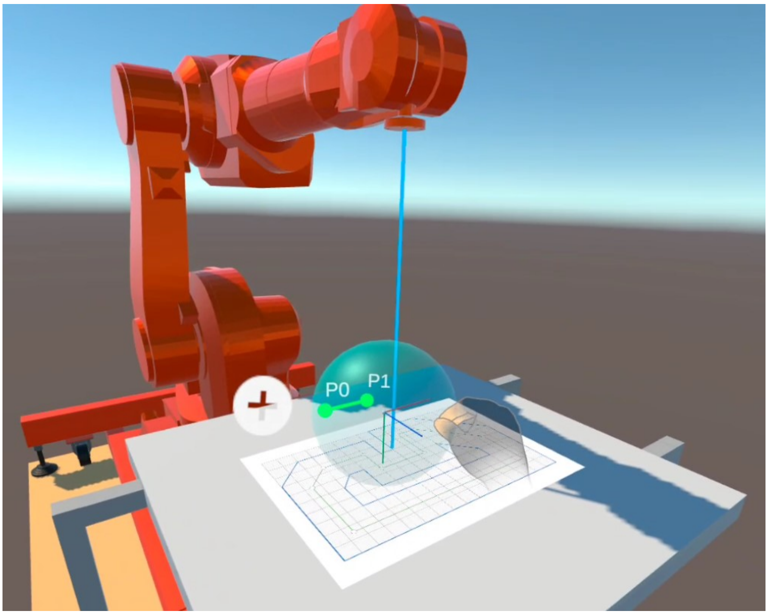

Figure 26: VR experimental setup with ABB IRB1600 path and path planning tool.

both robots. The ABBIRB1600 robot, corresponding MoveIt DT, running on the dedicated server, and AR setup are shown in Figure 27.

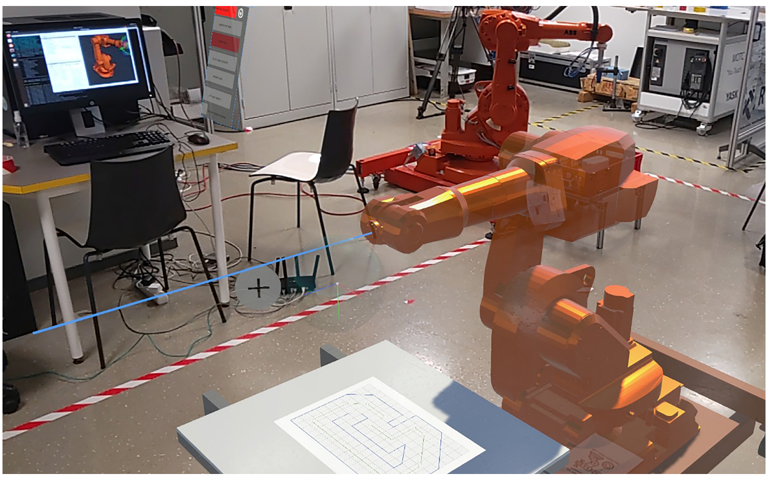
Figure 27: IRB1600 AR and real world setup

4.3.3 Experimental protocol

Two different sets of experiments took place with the same interfaces and two different robot models in the above mentioned university premises: IRB1600 in Tallinn University of Technology, and IRB1200 in the Technical University of the Shannon. A total of 42 subjects took place in the assessment. During the experiments users were asked to create ten path points in the positions visualized by the exemplifying path, as precise as possible and by using the provided interfaces and interactions methods.

Qualitative and quantitative data is collected during both sessions to be able to evaluate efficiency, effectiveness of the system, satisfaction and user experience while performing the tasks. These include total task time, heart rate, SUS , NasaTLX and UEQ questionnaires. A speak aloud protocol is adopted and users' comments recorded for further system improvements. Heart rate is not reported in the results as data was lost after the first round of 21 experiments due to a problem with the device hardware (Empatica4). The second session of experiments included path point coordinates for the AR, VR and real world planning tasks. The comparison of the collected path points with ground truth is addressed further in the text. The protocol has been reviewed and accepted by the Technical University of the Shannon ethics committee and is exemplified in Figure 28.

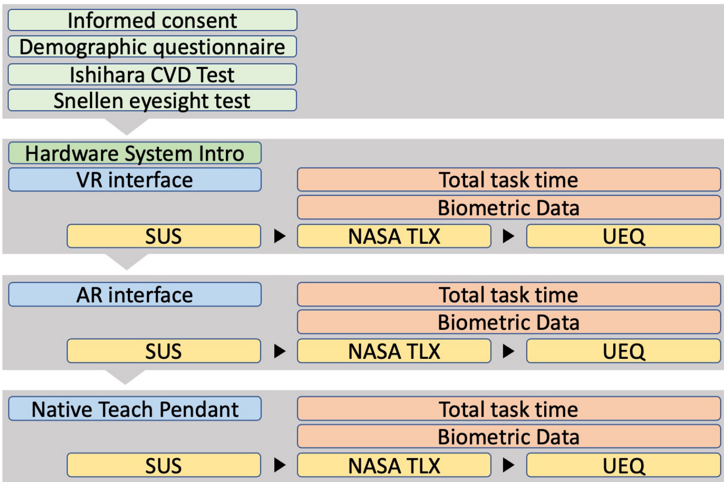

Figure 28: Experimental protocol

The use of the different interfaces was counterbalanced during both experimental sessions. An introductory session includes the explanation of study goals, and signing of an informed consent, an eye sight colour blindness test, the collection of demographic data, and a familiarization with the hardware and type of interaction methods.

4.3.4 Results of the first session

Demographics

The first session of experiments tested the proposed task and interface on 21 subjects at the Technical University of the Shannon with IRB1200 teach pendant. Table 14 reports the subject sample characteristics. Figure 29 shows the experimental setup with VR interface operated by the test subject side by side to the real robot.

A total of of 33 male and 8 female subject took part in the session with an average age of 33$\pm$ 5.8. Everyone had previous experience with VR headsets or equipment, while only one person did not have with AR. Only 5 people stated having experience with industrial robot programming and control. Three subjects out of the sample wore glasses and had eyesight below average. The preliminary questionnaire detected the expectations of the sample in interacting with XR technologies and industrial robots. These are scored on a 5 point Likert scale. Almost all the subjects agree on the potential XR technologies and interfaces with a resulting score of 4.6 (SD $\pm$0.5) in the case of VR and 4.5 (SD $\pm$0.6) in

Table 14: Characteristics of the sample

Age	33 ± 5.8
Gender	
Male	13
Female	8
Have experience with AR headsets	
Yes	20
No	1
Experience with VR headsets	
Yes	21
No	0
Experience with industrial robot programming	
Yes	5
No	16

case of AR. Regarding the expected enjoyability during use, results report an average 4.5 (SD $\pm$0.5) for both VR and AR. No subject had any concern, or felt scared, in having to operate an industrial robot (average 1.9, SD $\pm$0.9).

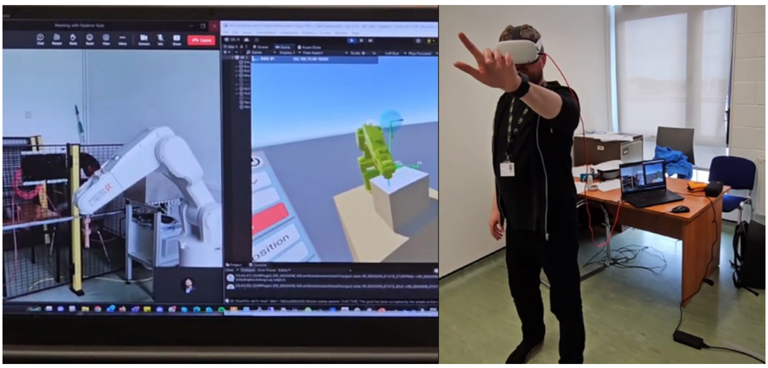

Figure 29: Experimental setup with IRC1200 operated by the user through the VR interface

Performance

Performance results confirm the VR UI is the most efficient interface to achieve the task followed by AR, while average timing to create the path took almost double for the native teach pendant. Table 15 and Figure 30 shows the results for the task performance (in seconds) for each type of interface.

Table 15: Task time performance, in seconds

	Mean	SD
Teach Pendant	334.24	$\pm$ 162.15
AR UI	179.57	$\pm$ 58.72
VR UI	152.76	$\pm$ 55.78

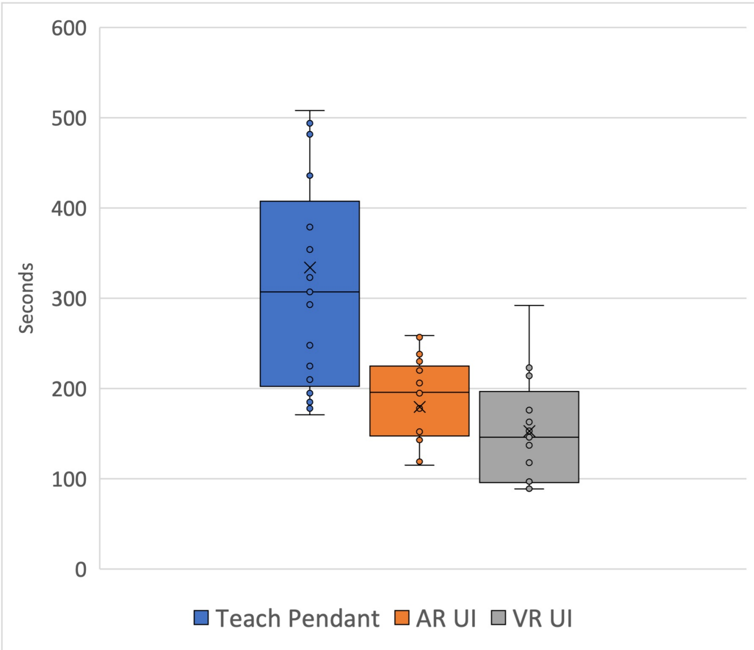

Figure 30: Task performance results

Task Load

Results of the NasaTLX questionnaire show higher results when using the native Tech Pendant, with a score of 43.89 overall. This is followed by AR UI (32.02) and VR UI (26.43) which seems to be the least demanding interface. Values for the temporal and physical demand are also above average in case of the Tech Pendant interface. In general perceived mental demand and effort for the three interfaces is above average. Table 16 shows the unweighted raw scores for the questionnaire while Figure 31 provides a visual representation of each questionnaire item.

Table 16: NasaTLX results

	AR UI	VR UI	Tech Pendant
Mental Demand	38.10	35.00	55.95
Physical Demand	26.43	21.90	45.95
Temporal Demand	29.75	28.10	44.29
Performance	33.00	25.24	33.33
Effort	35.95	28.57	48.81
Frustration	31.90	19.76	35.00
OVERALL	32.02	26.43	43.89

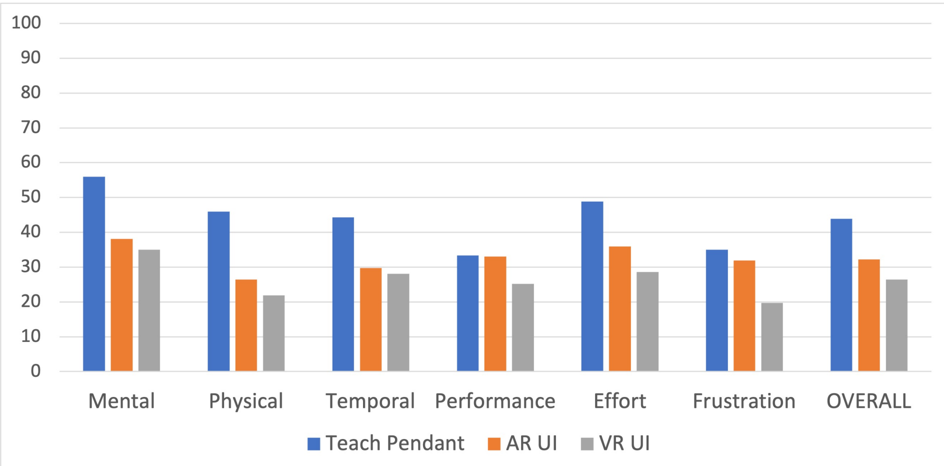

Figure 31: NasaTLX questionnaire results

Usability of the system and User Experience

UEQ results show a good user experience level for XR interfaces while only average for the Tech Pendant. Pragmatic quality perception for the native robot interface is low (0.51) as much as the perceived hedonic qualities (0.84). Both AR and VR interfaces score above average values for the hedonic quality with 2.28 and 2.21 respectively. VR has overall the highest scores in both categories (1.92 and 2.21). Results of the UEQ are shown in Table 17.

Results for the System Usability Scale questionnaire, Table 18 and Figure 32, show low level of usability for the Tech Pendant (Mean 50.83 SD ±24.71) corresponding to marginal or poor usability level. Usability for the AR UI is average resulting with a below good score of 69.76 (SD ±21.08). Better results are shown by the VR interface with a ranking of good or acceptable (76.54, SD ±20.47) for the specific interface. Overall high SD values confirm the high variability in the population and necessity for further experimental sessions.

Table 17: User Experience Questionnaire results

	Pragmatic	Hedonic	OVERALL
Teach Pendant	0.51	0.84	0.68
AR UI	1.55	2.28	1.92
VR UI	1.92	2.21	2.06

Table 18: System Usability Scale questionnaire results

	Mean	SD
Teach Pendant	50.83	± 24.71
AR UI	69.76	± 21.08
VR UI	76.54	± 20.47

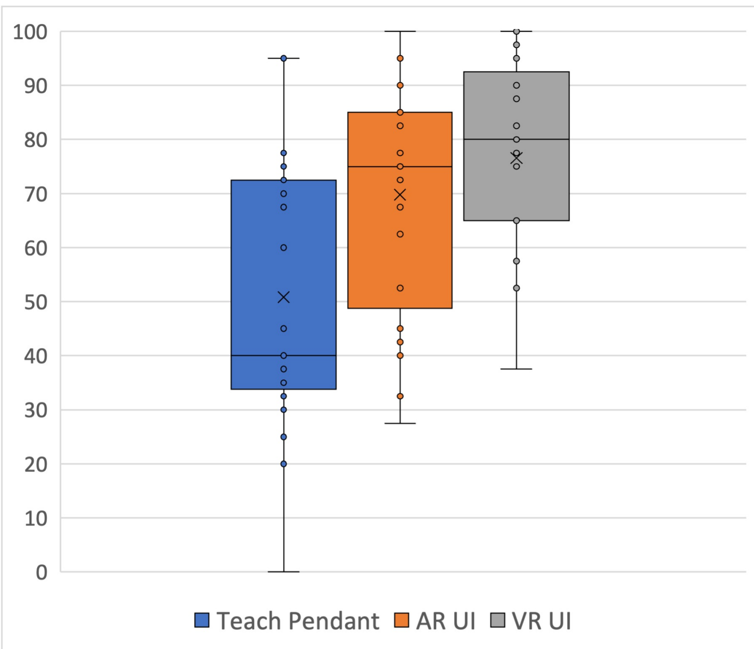

Figure 32: System Usability Scale boxplot

4.3.5 Results of the second session

Demographics

The second experimental session took place in Tallinn University of Technology adopting the same XR UIs and Tech Pendant model on a IRB1600 robot. A total of 21 subjects, 5 female and 16 male, took part in the experiments. Average age is 33 (±11.5) years old. 8 subject wore glasses and only two had below average eyesight. No subject was reported colour blind. Table 19 summarises the information about the subject sample for the second session of experiments. This group shows less expertise with AR and VR equipment with only 12 and 16 out of 21 subjects having already used the corresponding hardware. Also in case of industrial robot control and programming, only 9 subjects reported having experience with this technology. The introductory questionnaire, detecting expectations towards each technology, gave high scores for both AR and VR in terms of expected enjoyability, 4.7 (SD ±0.4) for AR and 4.5 (SD ±0.7) for VR. General interest towards the technology shows similar scores with a value of 4.7 (SD ±0.4) for both AR and VR. Also in this case nobody felt scared in interacting with an industrial robot (1.5 SD ±0.8).

Table 19: Characteristics of the sample, session 2

Age	33 ± 11.6
Gender	
Male	16
Female	5
Have experience with AR headsets	
Yes	12
No	9
Experience with VR headsets	
Yes	16
No	5
Experience with industrial robot programming	
Yes	9
No	12

Performance

Also in the second session of experiments performance results confirm the VR UI is the most efficient interface to achieve the task followed in this case by the Teach Pendant giving slightly better results in terms of timing than the AR interface. Table 20 and Figure 33 shows the results for the task performance (in seconds) for each type of interface.

Table 20: Task time performance (in seconds), session 2

	Mean	SD
Teach Pendant	255.86	$\pm$ 76.26
AR UI	259.29	$\pm$ 101.62
VR UI	206,24	$\pm$ 116.81

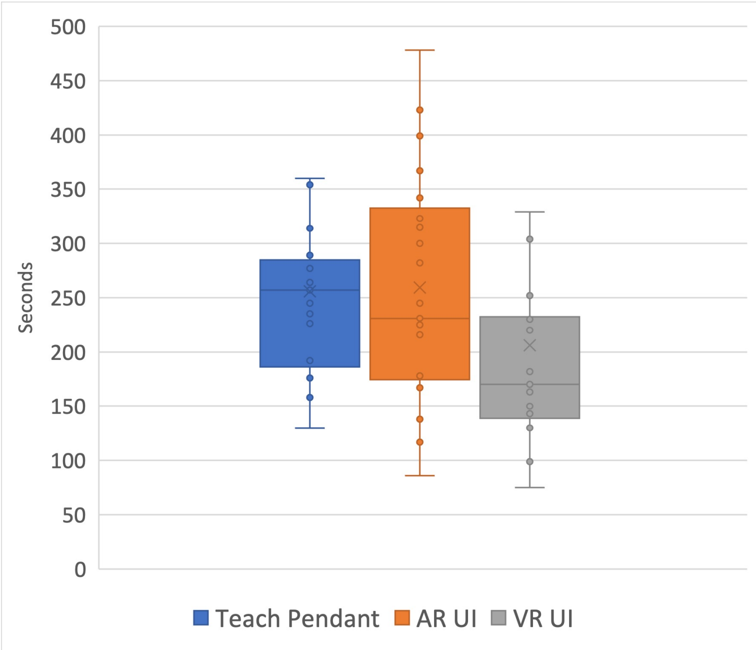

Figure 33: Task performance results, session 2

Task Load

Results of the NasaTLX questionnaire show higher results when using the native Tech Pendant, with a score of 45.20 overall, followed by AR UI (44.76) and VR UI (36.39). Table 21 shows the unweighted raw scores for the questionnaire while Figure 34 provides a visual representation of each questionnaire item.

Table 21: NasaTLX results, session 2

	AR UI	VR UI	Tech Pendant
Mental Demand	33.33	30.00	34.52
Physical Demand	38.1	25.48	37.38
Temporal Demand	42.86	29.05	39.52
Performance	64.76	70.48	73.81
Effort	43.33	34.52	49.29
Frustration	46.19	28.81	36.67
OVERALL	44.76	36.39	45.20

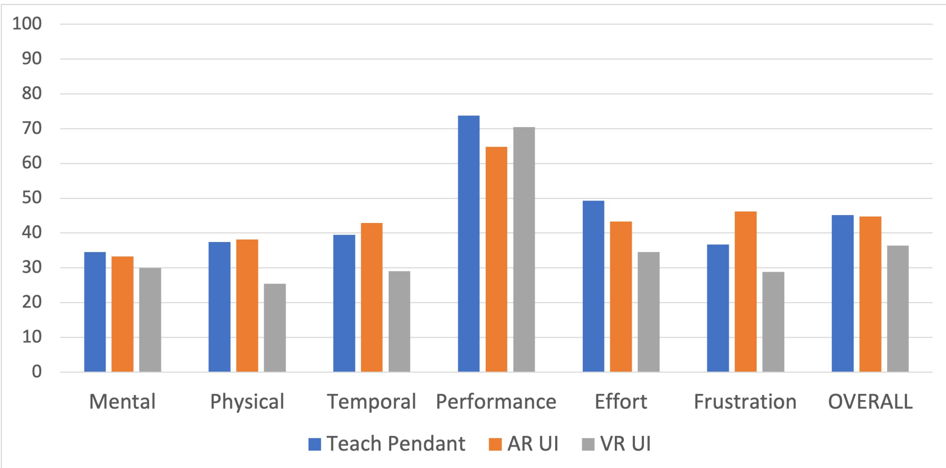

Figure 34: NasaTLX questionnaire results, session 2

Usability of the system and User Experience

UEQ reports a good user experience level for both AR (1.92) and VR (2.06) and low scores for the Teach Pendant interface (0.80). Pragmatic quality perception for the native robot interface is low (0.92) as much as the perceived hedonic qualities (0.65). As in the first session, AR and VR interfaces score above average values for the hedonic quality with 2.28 and 2.14 respectively. Results of the UEQ for the second session of experiments are overall in line with the results of the first session and shown in Table 22.

Results for the System Usability Scale questionnaire show low level of usability for the Tech Pendant (58.45 SD ±16.60) corresponding, again, to marginal usability level. Usability for the AR UI is also marginal with a score of 66.45 (SD ±19.46). The VR interface ranks good or acceptable usability level with a value of 79.28 (SD ±11.65). Results of the SUS questionnaire are shown in the following Table 23 and Figure 35.

Table 22: *User Experience Questionnaire results, session 2*

	Pragmatic	Hedonic	OVERALL
Teach Pendant	0.92	0.65	0.80
AR UI	1.15	2.28	1.72
VR UI	1.94	2.14	2.04

Table 23: System Usability Scale questionnaire results, session 2

	Mean	SD
Teach Pendant	58.45	$\pm$ 16.60
AR UI	66.45	$\pm$ 19.46
VR UI	79.28	$\pm$ 11.65

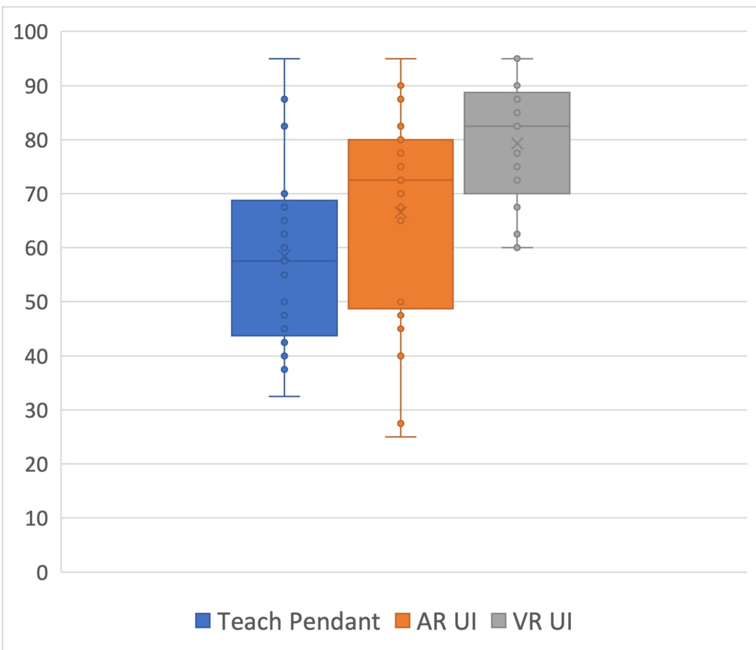

Figure 35: System Usability Scale boxplot, session 2

Evaluation of path planning accuracy

Path point coordinates where collected during the second experimental session to be able to compare the precision among planning tasks performed with different UIs. The ground truth for each point position was determined by calculating the relative distance from the robot base, constituting the origin of the axes for the physical robot path planning tasks. The position of the path points guiding the subjects during the experiments was, therefore, calculated based on this origin and subsequently transferred in the virtual and augmented reality environment. Path points coordinates are collected as JSON files in both the AR and VR UI applications. The systems collects information about point coordinate position and rotation. For the sake the experiment only positions along the X and Y axes are considered. Robot path point coordinates from the IRB1600 are downloaded from robot controller through a proprietary software. A custom parser is programmed to import the files relative to each user session with the three different interfaces, extract point coordinates and calculate the relative distance from the ground truth. Before the calculation a transformation of the AR and VR collected data points is made as Unity software uses a left-handed (Y-up) coordinate system. Table 24 reports the results of average point precision in millimeters (real-expected) on X and Y axes for each of the adopted UIs.

Table 24: Path point accuracy on X an Y axes (mm)

	X	SD	Y	SD	Overall
Teach Pendant	0.75	$\pm$0.44	1.19	$\pm$0.51	0.97 SD $\pm$0.52
AR UI	1.77	$\pm$1.49	2.16	$\pm$1.18	1.97 SD $\pm$1.36
VR UI	1.76	$\pm$1.43	2.1	$\pm$1	1.87 SD $\pm$1.25

Results show the higher accuracy of path planning task when operating the real robot. The overall difference from the ground truth is around 1mm for the teach pendant (0.97 SD$\pm$0.5) which is more than acceptable. Results for the AR and VR interfaces are similar with an overall 1.97 mm (SD $\pm$1.36) for the first and 1.87 (SD $\pm$1.25) for the second, resulting in a less accurate path planning task. Nevertheless, there is an high variability of the data from the mean, showing a possible low precision of the planning tool especially for the AR and VR setup.

Speak aloud protocol and systems errors

Several errors with the AR and VR systems interfaces were reported during both experimental sessions. The application running on Microsoft HoloLens had to be restarted a few times due to problem with image target tracking. In general the reflective and glossy floor surfaces resulted, in both cases, with several image target problems in case of the AR application. Due to this reason the results of path point precision for two of the subjects where not considered in the above mentioned comparison. 4 subjects saved path point twice or wrong points while programming the robot with the VR interface. This happened only once in case of the Teach Pendant sessions and four times in case of the AR interface. It was already stated that the impossibility of deleting specific points or editing their position is one of the major limitations of the proposed system.

A speak aloud protocol was used in both sessions. The main noticed problem was the sensitivity of the planning tool when 'grabbed' by the avatar hand. A problem of detecting the un-grab gesture (thumb and index finger opening) was in particular a major issue in a few sessions both while interacting with VR and AR UIs. One subject reported being annoyed by model edge flickering in the AR visualisation while another mentioned the brightness of the table stand in AR and VR being too high. A few users had problems interacting with the planning tool button to add path points when this was too close to the virtual stand. In general the higher position of the table stand in the VR environment was highly appreciated as it would facilitate the hand manipulation while standing and not leaning towards it. This was reported as an ergonomic problem of the AR UI setup. In general many users appreciated the hand manipulation option compared to the teach pendant joystick and capacitive screen. One subject proposed changing the planning tool overall appearance and make it more similar to a virtual pen.

4.3.6 Discussion

The two sessions of experiments confirm the higher perceived usability, better user experience and performance for the VR UI compared to native controls and AR. Results for the AR setups are non conclusive as they differ in the two sessions especially on the task performance side. This might be due to differing experience of the sample groups with AR and VR headsets and technologies. Further studies might considered a repeated measurement approach to evaluate the level of efficiency gain related to familiarity with the system. Generally, task load results seems to confirm the lower impact of the VR UI. Accuracy of AR and VR setups is largely comparable with native controls. Nevertheless, the goodness of these results have to be determined per use case and further investigated on different tasks. The speak aloud protocol provided very interesting feedback on the path planning tool design and the hand tracking system, together, of course, with general design of UI and visual feedback which can be integrated in future application developments.

5 Use cases in SME

This chapter introduces two pilot studies adopting the presented framework and XR based solutions aimed at the control of heavy machinery and the support of workers during assembly operations while improving design to production flow. The studies demonstrate the results from research and development projects of local SMEs and Tallinn University of Technology in the timber industry. This is one of the crucial sectors for local economy with a continuous growth and GDP impact in the past years [3] and many companies active along the wood harvesting, processing and product development.

5.1 XR based control and UI for a forestry crane (V)

The first work presents the implementation of a XR based interface for forestry crane control and operation visualization using HMD. The digitization of on field operations, harvesting and logging is already a reality [2], [107] with examples spanning from crane cabin operation support [121], XR based teleportation [164], or laser scanner crane posture estimation [54]. These applications are being developed to improve operators health and safety together with improving on field activities efficiency.

Hardware components and software architecture

The systems aims at modularity and scalability to different logging cranes and trailers produced by the local company BMF together with the use of different camera and controller systems. Based on the producer requirements the system needs to provide a real time immersive camera visualisation of the crane grapple and support legs together with an integrated view of crane hydraulic system parameters and a virtual replica of the real joysticks. One of the main goals is integrating the boom controls and UI navigation in the same logical layer to allow a smooth control from the same joystick pair. The separate logical layer for both controls and visualization is implemented in Unity. A separate Python software module works as a middleware between the hardware and controller and the unity application.

The hardware components of the proposed system include three cameras, a main one for the boom operations visualization, Teledyne Flir Oryx 10gE, and two secondary cameras for the trailer legs, Teledyne Flir Blackfly S GigE. The adopted immersive headset is an Oculus Quest 2. The system software is installed on a dedicated desktop PC which is intended to be located in the truck trailer compartment and powered by the truck battery, making the entire solution completely autonomous and portable. The machinery test-bench is produced by BMF. The crane (BMF650) [1] is installed on a trailer (BMF8T1E) provided by the same company, Figure 36. The controller (IQAN MC43), joysticks (IQAN LC5), Figure 37, and the controller interface (IQAN MD3) are produced by Parker.

The camera streams are handled by the Python application. Software language compatibility and native producers SDKs for Python guarantee the possibility of using different hardware without changing the Unity side of the system. The software modules is running on a server taking video input from the cameras and redirecting them to the Unity visualizer application through a TCP protocol. This receives the streams through a TCP socket and renders it on three different 3D panels in the immersive environment. To reduce latency the frames are sent in Bayer format. The de-bayer process happens in Unity before rendering the images as full-colour frames. The system architecture is shown in Figure 38. The server provides an interface to modify the camera settings from the visualizer application allowing the operator to set exposure or contrast within the immersive UI.

The crane is maneuvered by means of the joysticks and industrial controller taking care

Figure 36: BMF crane and trailer

Figure 37: The crane joysticks

of monitoring the state of the hardware components and control logic. The Unity software is not integrated in the control loop but works as a visualization application for the cameras and joystick positions. Similarly to the camera software system a Python module handles the connection between the crane controller and Unity. The controller supports CAN protocol with 5 CAN busses. The scaled percentile of the analogue voltage reads are exposed to one of the CAN bus. A CAN/USB or CAN/ETH adapter is used to forward the can frames to the dedicated PC. As much as for the cameras a TCP socket is used to con-

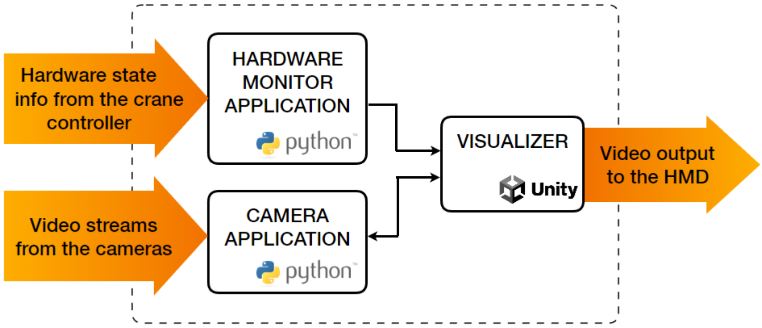

Figure 38: Software architecture of the system

nect Unity to the application, allowing, again, flexibility in exchanging hardware without having to modify the visualiser software. The crane control flow is shown in Figure 39.

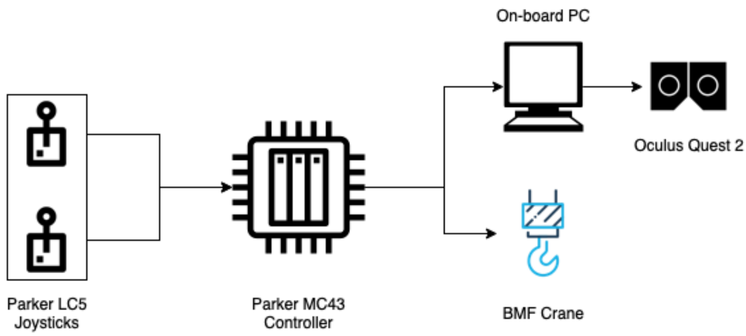

Figure 39: Crane control flow diagram

A combination of buttons on the native joysticks allows switching from crane control to UI navigation and option selection within the immersive UI. This feature prevents operators to start moving the crane while interacting, for instance, with the UI or camera visualization settings.

User Interface

The virtual UI integrates only some of the functions available in the native controller. The UI elements navigation and selection happens thought joystick selection and button press on both right and left joysticks. The Unity input mapping allows to connect different buttons and joystick positions to UI controls. The main UI visualises the real-time information about the crane boom extension, rotation, slewing and grapple opening. Other UI sections are dedicated to camera settings and adjustments, system info (such as date and time) Figure and show the main UI and the control mode of the immersive visualizer application developed in Unity.

Discussion

The main system functions have been tested locally, giving acceptable control and visualization results on the test bench crane provided by the company. The controller and joystick connection through CAN is established and stable. Limitations of the system, other than including more UI functionalities, are mainly related to testing with real operators,

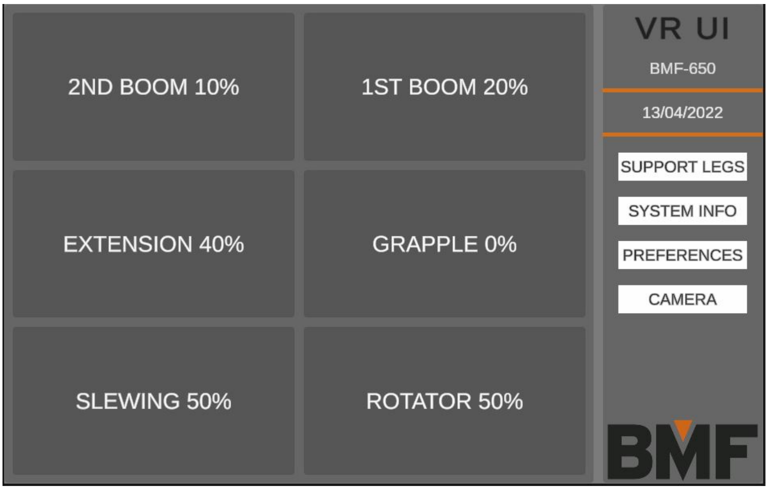

Figure 40: Immersive UI sections and controls

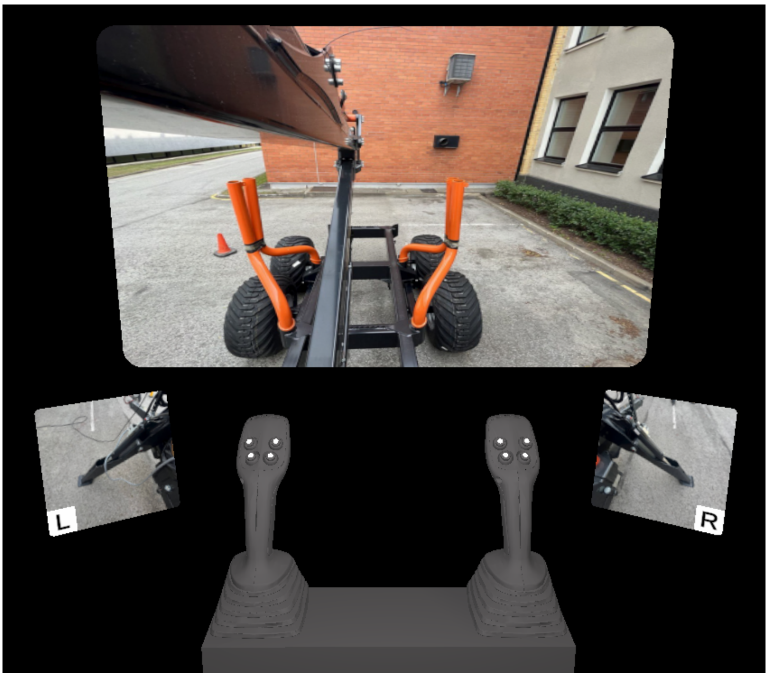

Figure 41: Crane control UI with joystick avatars and camera visualization panels

comparing traditional controls with the new system and providing feedback for further improvements of the UI and camera visualization. Camera based safety, and obstacle detection methods might be also considered in the future.

5.2 DT XR based user interface for operator assembly support (VI)

The second publication in this section presents a more complete example of adoption of UCD approach and implementations of XR based solutions in two Estonian SMEs active in the timber industry. The companies deal respectively with production of wood processing and packaging machinery and wooden house modules. The study is developed in the context of a European project aimed at the integration of 5G connectivity into SMEs to support the digitalization of the production process and adoption of advanced camera based or XR systems and services to improve efficiency and safety on the production floor. This scenario is quite relevant and consistent with the I4.0 and I5.0 technology and philosophy integration. As already mentioned this is not always easy [35] and presents several challenges still to be overcome, especially in regards to XR technologies adoption.

The general goals of the proposed applications are the support of machine servicing, referred as use case 1 (UC1), and assembly assistance to operator in the production floor including improvements in the design to production loop, use case 2 (UC2). Both systems should eventually rely on 5G connectivity and PLM based DT of the product for data retrieval and updates. The application development started with the acquisition of requirements from the factory management, design department, operators, and observations of the production floor operations. After a first round of mock-ups and interactable mock-ups definition, the integration of product data (e.g., 3D models, assembly or servicing instructions) continued towards a first application assessment within the owner company, only for UC2.

Requirements and KPIs

The UC1 refers to an XR application aimed at supporting operators and technicians in the servicing of wood processing and packaging machines produced by the owner company. The main idea is providing a servicing tool for each machine, capable of sourcing updated data from the scheduling software and provide contextualised AR based instructions. The identified requirements are the following:

- AR based instructions and content visualization.
- Image target or QR code localization on the specific machine.
- UI guidance on the required servicing tasks.

The UC2 application targets the product assembly procedures, the report of nonconformities (NC), and the overall improvements of the design to production loop thought the adoption PLM software. The required features for the application are the following:

- Operator login and profile management through QR code.
- Dynamic and animated 3D visualization of the component models and assembly steps.
- AR visualization of the component models.
- Assembly guidance information.
- Possibility of report nonconformities with different media (e.g., photo, audio or text)
- Flexibility of the application to run on different devices.

A few KPIs were selected to assess and measure, on a medium and longer term, the performance and the effectiveness of the adopted systems. The idea is demonstrating the benefits of the adoption of XR based operator support application rooted on DTs and local high speed connectivity. The selected key performance parameters are the following:

- Servicing and faults report time.
- Operator training time on a specific product assembly.
- Reduction of process lead time.
- Number of nonconformities.

Software architecture

Regarding UC1 the application is in an initial development stage. As the packing machine targeted for the project is installed in a third party company, and not always accessible, laboratory demonstrator application has been developed locally on a Flexible Manufacturing System. The application is developed using Unity game engine and Vuforia SDK for image/QR-code targeting and AR content visualization and alignment to real world coordinate system. The application is installed on a android based tablet.

For UC2 a different set of software is employed. The main reason is to respond to the company requirement regarding hardware use flexibility on the production floor (company owned tablets, new tablets, HMDs). The choice was therefore directed to web based tools with Next.js [4] being the main development framework, Tailwind.css [6] libraries used for the UI developement, and Three.js [7] for the integration of advanced 3D content visualization and interaction. The AR visualization feature of the application is made possible thought Three.js XR functionalities based on the well known WebXR APIs [10]. Differently from UC1 the placement of the AR content is based on ground detection properties. The main system architecture comprises of two parts shown in Figure 42.

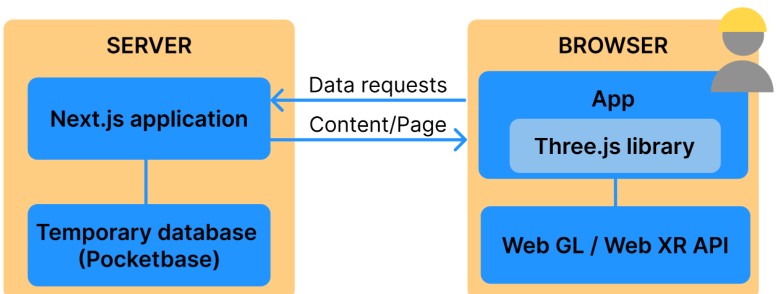

Figure 42: UC2 system architecture diagram

The first application is stored on a local server with a provisional database which in the future should be replaced by PLM. The Next.js based application sends the page templates to the web browser to allow content visualization and rendering through device hardware. The browser uses HTTPS protocol to fetch the required information from the API endpoints.

User Interface

Both use cases applications are optimised for tablet, therefore interactions with the digital content happens through tap, or pinch to zoom/rotate gestures.

Figure 43 shows the software information architecture for UC1. At this stage, information are saved locally. Each machine holds information on the scheduled servicing and related operations and descriptions.

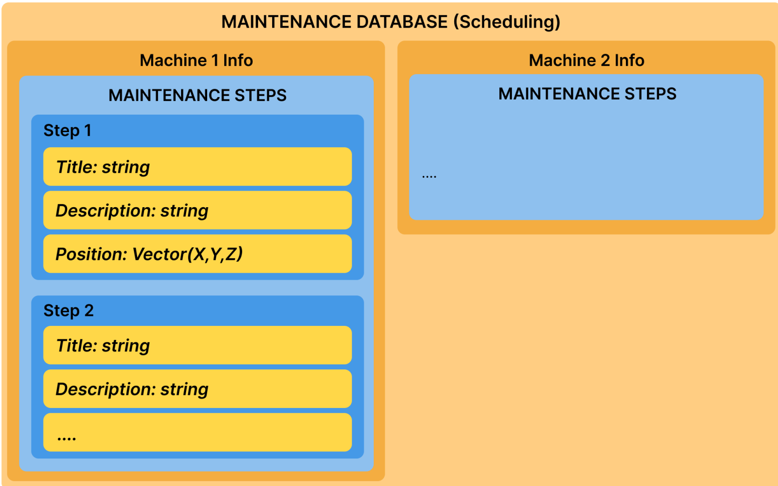

Figure 43: UC1 Information architecture diagram

The operator can visualise these as cards in AR overlaid on the real machine through QR code or image target tracking. The targeting happens thought the device camera. Figure 44 shows a picture in picture representation of a servicing operation. The smaller image shows the operator holding the tablet and targeting the QR code. The large image shows what the operator sees on the tablet monitor, in this case a dummy card with instructions and localisation of the task to be performed.

Figure 44: Picture in picture visualization of the maintenance application UI (UC1)

The UC2 application for product assembly support offers more functionalities and has more complex information architecture. This is shown in Figure 45.

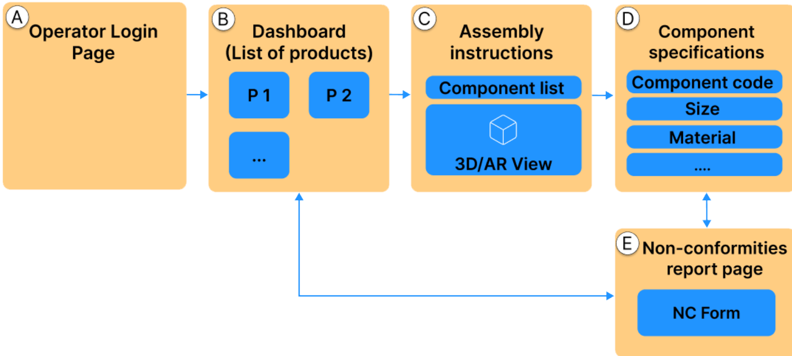

Figure 45: UC2 Information architecture diagram

The application is structured in different pages. The system login page (A) allows operators to access the application on a personal profile by scanning their employee card QR code. The project dashboard (B) provides an overview of the products and components to be assembled or in progress. These are shown as interactable 3D models, together with a status bar and notification about possible observed nonconformities. The assembly instruction page (C) provides a dynamic visualization of the assembly steps with information about each part and 3D model animations, Figure 46. As mentioned the operator can zoom and rotate models for better visualization. From this page it is also possible to access the AR functionality and place the component in real world space as shown in Figure 47.

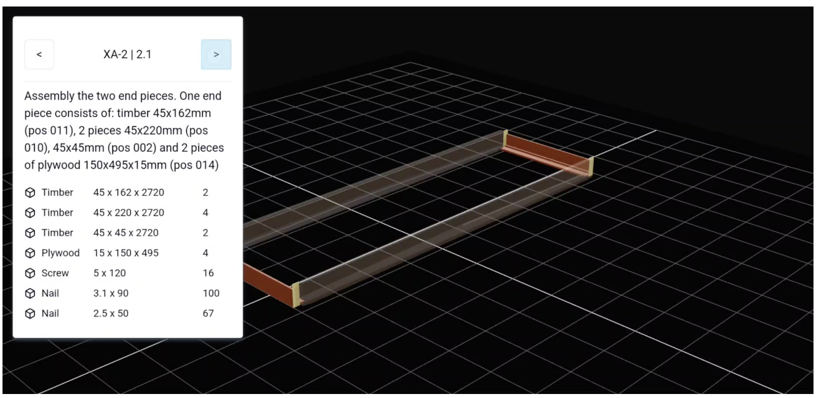

Figure 46: Assembly step page (C)

By clicking on each of the parts listed in the assembly instruction guidance page (C) a specification pop-up window (D) is visualized providing more detailed information on the part ID code, size, type of material etc. From this page it is also possible to access the non conformity report form (E), Figure 48.

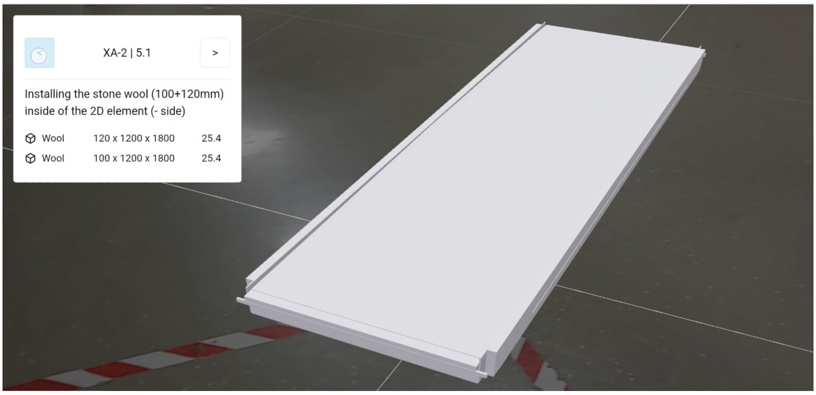

Figure 47: Assembly step page with AR visualization

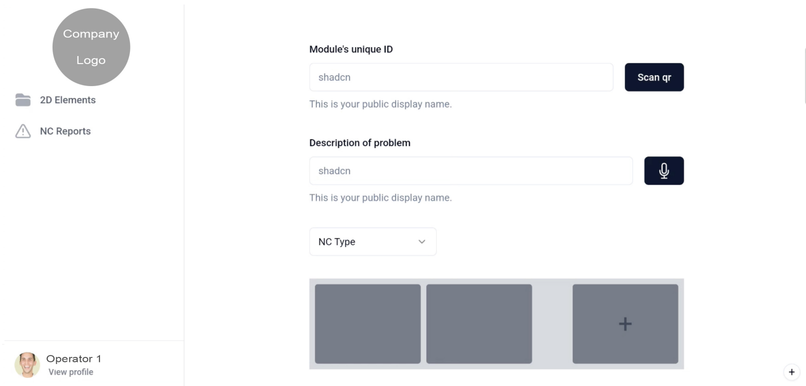

Figure 48: Non-conformity report page

This functionality follows the structure of the existing company NC software. The page allows selecting among different standard NCs and adding annotations in form of text, videos or images.

Assessment of the system (UC2)

The assembly support application (UC2) have been preliminary tested in the company premises with a group of workers from different departments. The experimental protocol consisted in two different UI navigation and interaction tasks. After a briefing on the study goals, tasks instructions, risks, benefits and the signature of an informed consent, the subjects were asked to perform an intuitive navigation of the UI, first, and to report two nonconformities of two different product components. After tasks completion the subjects completed three different subjective questionnaires to assess user experience and usability of the application. For usability and UX evaluation the SUS and UEQ questionnaires are employed. A custom questionnaire collected basic demographics data and a subjective response on UI elements structure and accessibility. These are scored on a 5 point Likert scale with a positive and negative term for each proposed statement. During the test we asked users to freely comment on the application functions and navigation and report their thoughts on the overall usability and efficiency of the solution.

The sample consisted in a group of employees from different departments of the owner company including two factory workers, one employee in pre-production one production department manager, one employee in the IT and one in the management department. The average age of the sample is 35 year old (SD $\pm$7). Results from the SUS questionnaire gave good usability results with an average of 72.9 (SD $\pm$9.7) corresponding to acceptable. UEQ gave a good response for the pragmatic qualities with a scoring of 2.54 and overall UX scoring of 2.1. Hedonic qualities have a lower but still good evaluation with a scoring of 1.67. Results for the custom questionnaire are shown in Table 25.

Table 25: Custom questionnaire results

	Mean	SD
Interface buttons are clearly visible	4.3	$\pm$0.7
The application controls are clear and accessible	4.3	$\pm$0.5
The component 3D visualisation is intuitive and clear	4.8	$\pm$0.4
The navigation of the interface is complicated	3.5	$\pm$0.9
I could not find the pages I needed to access	2.3	$\pm$1.6
Functions I needed are easy to find	3.8	$\pm$0.6

Overall results of the questionnaire give a good feedback on the visibility and clearness of the UI elements and functions. The speak aloud protocol revealed some problems with page navigation, especially between additional documentation, assembly steps and dashboard. A too high sensitivity of the 3D model navigation (zoom and pan) was reported by one operator. More than one subject agreed on the necessity of having some guidance in the activation of the AR view and breadcrumbs, especially in the assembly steps procedure. The overall development and assessment process, despite being limited by the lack of proper PLM software setup, and bound to a small test group, can be considered successful. Improvements based on the feedback are already ongoing together with further developments in the data management flow.

5.2.1 Discussion

The work presents a successful example of iterative (UCD and agile) approach to the implementation of XR based tools and application in two different companies active in the Estonian timber industry. The preliminary results on a first release of the application supporting operators in the assembly of wooden house modules has given good results in terms of usability and user experience. Furthermore, the applications are scalable to include other products and functions, constituting a base for service oriented product development for both companies. Further development, PLM repository integration, and medium and long term assessment based on the defined KPIs is necessary to establish the efficiency and effectiveness of the solutions and refined their integration in the design to production loop.

6 Conclusions

All research questions presented in section 2 have been addressed in the corresponding thesis, with specific references indicating which chapter and paper addressed each question, as detailed in Paragraph 2.4. The subsequent paragraphs will provide a reflective discussion on the findings, 6.1, identify limitations and propose future research directions, 6.2, and present feedback from both research and industrial domains in paragraph 6.3.

6.1 Discussion

This work tackles different aspects of the adoption of advanced XR interfaces within the context of I5.0 and HRI. DTs and XR become the core of the proposed framework aimed at the implementation of I5.0 objectives. DTs and UCD loops integration are leveraged together with XR technologies immersive interaction and visualisation potential for the inclusion of workers within the cyber-physical industrial domain. (I publication).

The efficacy and centrality of the proposed framework, grounding on DT based XR interfaces (RQ1, I, II, III, IV, V, VI publication) has been demonstrated both in laboratory and real world use cases. DT-based XR interfaces provide a core method to address I5.0 needs and requirements. The approach is shifting the focus of development and evaluation towards operators, with positive impacts observed across various cultural backgrounds and scales of SMEs. Furthermore, the practical implementation of these methods has been tested in local companies, confirming their real-world applicability and effectiveness in enhancing user experience and operational efficiency.

XR interfaces demonstrated reliability, efficiency and effectiveness, (II, III, and IV publication) towards lab based use cases and real world assembly tasks in industry (RQ2). XR UIs are largely comparable to real world controls for industrial robot operation with a trend of lower task load on the users, higher efficiency, especially for HMD based VR solutions, and comparable precision on a path planning tasks.

A set of qualitative and quantitative objective and subjective metrics and methods, defined along the UCD approach and sourcing from DT based XR technologies, have been demonstrated as valuable for the assessment of these systems both in laboratory and real world evaluations (RQ3, II, III, IV, VI publication). These tests gave the opportunity to refine the experimental protocols, limiting unsuitable or unripe methods and metrics, and detecting system limitations which should be addressed in the future.

As already mentioned the UCD approach together with the proposed interaction and visualization methods and tools have been successfully, although partially, implemented and assessed in real world use cases with local SMEs. XR has been adopted as a support tool for machine operation, servicing, maintenance and operator training (RQ4, V and VI publication). XR UIs provide a way to improve user experience, comfort and safety, while enhancing the efficiency of the design to production loop through integration of advanced data visualization (not necessarily immersive) and interactions.

6.2 Limitations and future work

Limitations and necessary improvements to the proposed work are manifold. Firstly, one of the main limitation of this, and other cited studies, is the evident lack of multidisciplinary approach. The topics under discussion encompass subjects belonging to engineering, design, information technology, ergonomics, psychology, sociology if not the medical field. Regardless of the industrial, HRI and I5.0 context of these applications the approach needs to be multifaceted and multidisciplinary, which seems to be, in most of the cases, a hard task to achieve both in academics and industry.

Secondly, the assessment plan (Figure 8, section 3) should be completed to integrate ergonomics evaluation of different type of HMDs or hand held equipment, on the computer system interaction side, and DT synchronization and visualization, on the other. A more complete study plan should encompass different types of XR interaction methods and a more accurate evaluation of newer hand tracking systems, tantamount publication II experiments, along with different HMDs equipment. A part of this further study plan could focus on different interaction metaphors beyond the real world physics, especially connected to industrial systems collaborative methods or user machine embodiment in the XR domain through posture tracking.

A third point would be the evaluation, extension and standardization of assessment practices, including tools and equipment adopted for the purpose. As initial results are non conclusive, there is a need to investigate the efficiency and benefits of wearable biofeedback devices and test them in real world scenarios. More reliable methods for eye and attention tracking have to be tested end implemented both in XR and physical setups.

A standardised and open set of tools, especially in the academic domain, would enable an homogeneous and repeatable scenario of user studies across different institutions and locations. A more in depth statistical analysis should be consider correlating factors such as age or experience with both industrial and XR equipment on a larger and more diverse experimental group, across different countries and cultural backgrounds.

A thought should be given to the dynamic scenario of hardware and software developments in the XR domain. This comes along with both uncertainties on long term support and reliability of the systems, which can be partially overcome by open source, cross platform and hardware agnostic software packages, and an increasingly efficient and groundbreaking quality of natural interactions (hand tracking, voice) driven by sensor and AI integration. The latter is increasingly challenging the ways we interact with digital contents therefore in need of continuously updated assessment tools and methods.

Furthermore the need of investigating the medium and long term benefits for such technologies in SMEs is paramount. Regardless of the numerous applications in training, maintenance and support, the uncertainties on the added value and company revenues from XR software and hardware applications is still high. Moreover, decentralised and "on demand" industrial scenarios are currently developing and might benefit from the remote assistance and control supported by XR in the near future. More focus should be given to the assessment of the psychological, social and ethics impact of these technologies on workers and work organisation, together with clear guidelines and boundaries on human based data sourcing and privacy.

The proposed approach is, nevertheless, an iterative, open and non conclusive one. These can be considered as as set of guidelines and good practices based on repetitive and validated experimental sessions, which can and must be stretched, improved and adjusted on the go, depending on specific requirements and needs, while keeping the same roots and structure.

6.3 Impact

The results of the presented studies and use cases, together with the suggested research plan along the UCD loop, can be both beneficial to create a common ground for laboratory experiments and an integrated (XR, UCD, DT) approach to real world application developments with human operators as main protagonists: the Operator4.0 and human cyber-physical systems mentioned in section 1. This is answering the needs and expectations of today's industrial developments, for flexible, resilient, decentralised and user-centric systems in I5.0, providing a structured solution on how different realms might converge on XR as a main interface and medium for the operators towards the digital domain. The proposed process could boost the adoption and assessment of DT XR based technologies in SMEs by providing valuable and validated procedures together with tools to source from, without starting from scratch.

Moreover, beyond just the applications and examples presented in this work, the author recognizes a broader potential for the findings in other domains. These include the optimization of workspace and assembly lines layout, and providing assembly guidelines in different fields, for instance, manual assembly, hazardous environments, such as mines, or control applications in the space and marine fields.

Insights gained from user studies can address issues related to continuous education and aging workforce accessibility, enabling individuals to engage with simpler interfaces for enhanced learning and prolonged efficiency in labor. There is also a clear synergy between these technologies and the healthcare domain with undeniable potential for shared insight on human well-being and how XR technologies could support everyday activities.

The framework could eventually be undertaken as a general cross-institutional plan to deepen the knowledge and refine the tools to safely and efficiently integrate humans and machines in collaborative systems, and move forward in the definition of standards and metrics for XR based HRI.

References

[1] BMF forestry trailers & cranes. https://www.bmf.ee/products/cranes/bmf-650, (Accessed on April 10 2024).

[2] Digital revolution in the Estonian forestry and wood industry. https://estoniantimber.ee/best-practices/digital-revolution-in-the-estonian-forestry-and-wood-i/, (Accessed on April 10 2024).

[3] Estonian Ministry of Climate, Forestry. https://envir.ee/en/water-forest-resources/forestry, (Accessed on April 10 2024).

[4] Next.js. https://nextjs.org/, (Accessed on April 10 2024).

[5] Ros-industrial official website. https://rosindustrial.org/, (Accessed on April 10 2024).

[6] Tailwind.css. https://tailwindcss.com/, (Accessed on April 10 2024).

[7] Three.js. https://threejs.org/, (Accessed on April 10 2024).

[8] Unity game engine. https://unity.com/, (Accessed on April 10 2024).

[9] Virtual Experience Research accelerator (VERA), University of Central Florida. https://sreal.ucf.edu/vera/, (Accessed on April 10 2024).

[10] Webxr. https://immersiveweb.dev/, (Accessed on April 10 2024).

[11] Xr and its potential for europe. Brussels, April 2021. https://xreuropepotential.com/assets/pdf/ecorys-xr-2021-report.pdf, (accessed on April 10 2024).

[12] M. Ali and R. E. Cardona-Rivera. Comparing gamepad and naturally-mapped controller effects on perceived virtual reality experiences. In *ACM Symposium on Applied Perception 2020*, pages 1–10, 2020.

[13] C. Anthes, R. J. García-Hernández, M. Wiedemann, and D. Kranzlmüller. State of the art of virtual reality technology. In *2016 IEEE aerospace conference*, pages 1–19. IEEE, 2016.

[14] B. D. Argall, S. Chernova, M. Veloso, and B. Browning. A survey of robot learning from demonstration. *Robotics and autonomous systems*, 57(5):469–483, 2009.

[15] Z. Bagdasarov, A. A. Martin, and M. R. Buckley. Working with robots: Organizational considerations. *Organizational dynamics*, 49(2):100679, 2020.

[16] B. R. Barricelli, E. Casiraghi, and D. Fogli. A survey on digital twin: Definitions, characteristics, applications, and design implications. *IEEE Access*, 7:167653–167671, 2019.

[17] C. Bartneck, D. Kulic, and E. A. Croft. Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots. In *Metrics for Human-Robot Interaction, ACM/IEEE International Conference on Human-Robot Interaction*, Amsterdam, The Netherlands, 2008. ACM.

[18] J. Berg and S. Lu. Review of interfaces for industrial human-robot interaction. *Current Robotics Reports*, 1:27–34, 2020.

[19] L. P. Berg and J. M. Vance. Industry use of virtual reality in product design and manufacturing: a survey. *Virtual reality*, 21:1–17, 2017.

[20] R. A. Bolt. "Put-that-there" Voice and gesture at the graphics interface. In *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*, pages 262–270, 1980.

[21] Y. Bondarenko, V. Kuts, S. Pizzagalli, K. Nutonen, N. Murray, and E. O'Connell. Universal xr framework architecture based on open-source xr tools. In *International XR Conference*, pages 87–98. Springer, 2023.

[22] Y. Bondarenko, S. L. Pizzagalli, V. Kuts, E. Petlenkov, and T. Otto. Digital twin based learning from demonstration system for industrial robots. In *ASME International Mechanical Engineering Congress and Exposition*, volume 87608, page V003T03A081. American Society of Mechanical Engineers, 2023.

[23] J. Brooke et al. SUS-A Quick and dirty Usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[24] K. Bruynseels, F. Santoni de Sio, and J. Van den Hoven. Digital twins in health care: ethical implications of an emerging engineering paradigm. *Frontiers in genetics*, 9:31, 2018.

[25] T. Buldakova and S. Suyatinov. Hierarchy of human operator models for digital twin. In *2019 International Russian Automation Conference (RusAutoCon)*, pages 1–5. IEEE, 2019.

[26] G. C. Burdea and P. Coiffet. *Virtual reality technology*. John Wiley & Sons, 2003.

[27] A. Burghardt, D. Szybicki, P. Gierlak, K. Kurc, P. Pietruś, and R. Cygan. Programming of industrial robots using virtual reality and digital twins. *Applied Sciences*, 10(2):486, 2020.

[28] J. R. Campana and M. Quaresma. The importance of specific usability guidelines for robot user interfaces. In *Design, User Experience, and Usability: Designing Pleasurable Experiences: 6th International Conference, DUXU 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part II 6*, pages 471–483. Springer, 2017.

[29] S. Chitta. MoveIt!: An introduction. In *Studies in Computational Intelligence*, pages 3–27. Springer International Publishing, 2016.

[30] V. Clay, P. König, and S. Koenig. Eye tracking in virtual reality. *Journal of eye movement research*, 12(1), 2019.

[31] E. Commission, D.-G. for Research, Innovation, M. Breque, L. De Nul, and A. Petridis. *Industry 5.0 – Towards a sustainable, human-centric and resilient European industry*. Publications Office of the European Union, 2021.

[32] D. Concannon, R. Flynn, and N. Murray. A quality of experience evaluation system and research challenges for networked virtual reality-based teleoperation applications. In *Proceedings of the 11th ACM workshop on immersive mixed and virtual environment systems*, pages 10–12, 2019.

[33] L. Damiani, M. Demartini, G. Guizzi, R. Revetria, and F. Tonelli. Augmented and virtual reality applications in industrial systems: A qualitative review towards the industry 4.0 era. *IFAC-PapersOnLine*, 51(11):624–630, 2018.

[34] A. De Luca and F. Flacco. Integrated control for phri: Collision avoidance, detection, reaction and collaboration. In *2012 4th IEEE RAS & EMBS international conference on biomedical robotics and biomechatronics (BioRob)*, pages 288–295. IEEE, 2012.

[35] L. F. de Souza Cardoso, F. C. M. Q. Mariano, and E. R. Zorzal. A survey of industrial augmented reality. *Computers & Industrial Engineering*, 139:106159, 2020.

[36] J. DelPreto, J. I. Lipton, L. Sanneman, A. J. Fay, C. Fourie, C. Choi, and D. Rus. Helping robots learn: a human-robot master-apprentice model using demonstrations via virtual reality teleoperation. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 10226–10233. IEEE, 2020.

[37] K. A. Demir, G. Döven, and B. Sezen. Industry 5.0 and human-robot co-working. *Procedia computer science*, 158:688–695, 2019.

[38] D. Diachenko, A. Partyshev, S. L. Pizzagalli, Y. Bondarenko, T. Otto, and V. Kuts. Industrial collaborative robot digital twin integration and control using robot operating system. *Journal of Machine Engineering*, 22, 2022.

[39] M. Dianatfar, J. Latokartano, and M. Lanz. Review on existing vr/ar solutions in human–robot collaboration. *Procedia CIRP*, 97:407–411, 2021.

[40] S. Doolani, C. Wessels, V. Kanal, C. Sevastopoulos, A. Jaiswal, H. Nambiappan, and F. Makedon. A review of extended reality (xr) technologies for manufacturing training. *Technologies*, 8(4), 2020.

[41] B. Dumas, D. Lalanne, and S. Oviatt. Multimodal interfaces: A survey of principles, models and frameworks. In *Human machine interaction: Research results of the mmi program*, pages 3–26. Springer, 2009.

[42] J. Egger and T. Masood. Augmented reality in support of intelligent manufacturing–a systematic literature review. *Computers & Industrial Engineering*, 140:106195, 2020.

[43] U. Erra, D. Malandrino, and L. Pepe. Virtual reality interfaces for interacting with three-dimensional graphs. *International Journal of Human–Computer Interaction*, 35(1):75–88, 2019.

[44] M. Eswaran, A. K. Gulivindala, A. K. Inkulu, and M. R. Bahubalendruni. Augmented reality-based guidance in product assembly and maintenance/repair perspective: A state of the art review on challenges and opportunities. *Expert Systems with Applications*, 213:118983, 2023.

[45] L. Gong, Å. Fast-Berglund, and B. Johansson. A framework for extended reality system development in manufacturing. *IEEE Access*, 9:24796–24813, 2021.

[46] I. Graessler and A. Pöhler. Integration of a digital twin as human representation in a scheduling procedure of a cyber-physical production system. In *2017 IEEE international conference on industrial engineering and engineering management (IEEM)*, pages 289–293. IEEE, 2017.

[47] M. Grieves. Digital twin: manufacturing excellence through virtual factory replication. *White paper*, 1(2014):1–7, 2014.

[48] C. Guo and E. Sharlin. Utilizing physical objects and metaphors for human robot interaction. In *AISB 2008 Convention Communication, Interaction and Social Intelligence*, volume 1, page 12. Citeseer, 2008.

[49] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[50] V. Havard, B. Jeanne, M. Lacomblez, and D. Baudry. Digital twin and virtual reality: a co-simulation environment for design and assessment of industrial workstations. *Production & Manufacturing Research*, 7(1):472–489, 2019.

[51] E. Helms, R. D. Schraft, and M. Hagele. rob@ work: Robot assistant in industrial environments. In *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication*, pages 399–404. IEEE, 2002.

[52] R. Hetrick, N. Amerson, B. Kim, E. Rosen, E. J. de Visser, and E. Phillips. Comparing virtual reality interfaces for the teleoperation of robots. In *2020 Systems and Information Engineering Design Symposium (SIEDS)*, pages 1–7. IEEE, 2020.

[53] G. Horváth and G. Erdős. Gesture control of cyber physical systems. *Procedia Cirp*, 63:184–188, 2017.

[54] H. Hyyti, V. V. Lehtola, and A. Visala. Forestry crane posture estimation with a two-dimensional laser scanner. *Journal of Field Robotics*, 35(7):1025–1049, 2018.

[55] International Organization for Standardization. ISO 25010 - Systems and software engineering - Systems and software quality requirements and evaluation (square) - System and software quality models. Standard, International Organization for Standardization, 2011.

[56] International Organization for Standardization. ISO/TS 15066:2016. Robots and robotic devices – Collaborative robots. Standard, International Organization for Standardization, 2016.

[57] International Organization for Standardization. ISO 9241-11. Ergonomics of human-system interaction. Part 11: Usability: Definitions and concepts. Standard, International Organization for Standardization, 2018.

[58] International Organization for Standardization. ISO 9241-210:2019. Ergonomics of Human-System Interaction. Human-Centred Design for Interactive Systems. Standard, International Organization for Standardization, 2019.

[59] A. Jaimes and N. Sebe. Multimodal human–computer interaction: A survey. *Computer vision and image understanding*, 108(1-2):116–134, 2007.

[60] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks. Characterising the digital twin: A systematic literature review. *CIRP journal of manufacturing science and technology*, 29:36–52, 2020.

[61] H. Kagermann, W. Wahlster, J. Helbig, et al. Recommendations for implementing the strategic initiative industrie 4.0: Final report of the industrie 4.0 working group. *Forschungsunion: Berlin, Germany*, 2013.

[62] K. Kalda, S.-L. Pizzagalli, R.-M. Soe, R. Sell, and M. Bellone. Language of driving for autonomous vehicles. *Applied Sciences*, 12(11):5406, 2022.

[63] C. Karagiannidis, A. Koumpis, and C. Stephanidis. Supporting adaptivity in intelligent user interfaces: The case of media and modalities allocation. In *Proceedings of the 1st ERCIM Workshop on User Interface for All*. Citeseer, 1995.

[64] A. Karakra, F. Fontanili, E. Lamine, J. Lamothe, and A. Taweel. Pervasive computing integrated discrete event simulation for a hospital digital twin. In *2018 IEEE/ACS 15th international conference on computer systems and Applications (AICCSA)*, pages 1–6. IEEE, 2018.

[65] W. Karwowski. The ocra method: Assessment of exposure to occupational repetitive actions of the upper limbs. In *International Encyclopedia of Ergonomics and Human Factors-3 Volume Set*, pages 3337–3345. CRC Press, 2006.

[66] N. Kayhani, H. Taghaddos, M. Noghabaee, et al. Utilization of virtual reality visualizations on heavy mobile crane planning for modular construction. *arXiv preprint arXiv:1901.06248*, 2019.

[67] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

[68] H. Kharoub, M. Lataifeh, and N. Ahmed. 3d user interface design and usability for immersive vr. *Applied sciences*, 9(22):4861, 2019.

[69] H. Kim and Y. Choi. Performance comparison of user interface devices for controlling mining software in virtual reality environments. *Applied Sciences*, 9(13):2584, 2019.

[70] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn. Digital twin in manufacturing: A categorical literature review and classification. *Ifac-PapersOnline*, 51(11):1016–1022, 2018.

[71] D. Krupke, F. Steinicke, P. Lubos, Y. Jonetzko, M. Görner, and J. Zhang. Comparison of multimodal heading and pointing gestures for co-located mixed reality human-robot interaction. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1–9, Madrid, Spain, 2018. IEEE, IEEE.

[72] V. Kuts, Y. Bondarenko, M. Gavriljuk, A. Paryshev, S. Jegorov, S. Pizzagall, and T. Otto. Digital twin: Universal user interface for online management of the manufacturing system. In *ASME International Mechanical Engineering Congress and Exposition*, volume 85567, page V02BT02A003. American Society of Mechanical Engineers, 2021.

[73] V. Kuts, J. A. Marvel, M. Aksu, **S.L. Pizzagalli**, M. Sarkans, Y. Bondarenko, and T. Otto. Digital twin as industrial robots manipulation validation tool. *Robotics*, 11(5):113, 2022.

[74] V. Kuts, T. Otto, T. Tähemaa, and Y. Bondarenko. Digital twin based synchronised control and simulation of the industrial robotic cell using virtual reality. *Journal of Machine Engineering*, 19(1):128–144, 2019.

[75] V. Kuts, M. Sarkans, T. Otto, T. Tähemaa, and Y. Bondarenko. Digital twin: Concept of hybrid programming for industrial robots—use case. In *ASME International Mechanical Engineering Congress and Exposition*, volume 59384, page V02BT02A005. American Society of Mechanical Engineers, 2019.

[76] H. Laaki, Y. Miche, and K. Tammi. Prototyping a digital twin for real time remote control over mobile networks: Application of remote surgery. *IEEE Access*, 7:20325–20336, 2019.

[77] E. Lampen, J. Teuber, F. Gaisbauer, T. Bär, T. Pfeiffer, and S. Wachsmuth. Combining simulation and augmented reality methods for enhanced worker assistance in manual assembly. *Procedia Cirp*, 81:588–593, 2019.

[78] B. Laugwitz, T. Held, and M. Schrepp. Construction and evaluation of a user experience questionnaire. *HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, November 20-21, 2008. Proceedings 4*, pages 63–76, 2008.

[79] D. Lee and Y. S. Park. Implementation of augmented teleoperation system based on robot operating system (ros). In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5497–5502, 2018.

[80] M. Liu, S. Fang, H. Dong, and C. Xu. Review of digital twin about concepts, technologies, and industrial applications. *Journal of Manufacturing Systems*, 2020.

[81] C. Lo, C. Chen, and R. Y. Zhong. A review of digital twin in product design and development. *Advanced Engineering Informatics*, 48:101297, 2021.

[82] F. Longo, A. Padovano, and S. Umbrello. Value-oriented and ethical technology engineering in industry 5.0: A human-centric perspective for the design of the factory of the future. *Applied Sciences*, 10(12):4182, 2020.

[83] P. K. R. Maddikunta, Q.-V. Pham, B. Prabadevi, N. Deepa, K. Dev, T. R. Gadekallu, R. Ruby, and M. Liyanage. Industry 5.0: A survey on enabling technologies and potential applications. *Journal of Industrial Information Integration*, 26:100257, 2022.

[84] K. Mahmood, **S.L. Pizzagalli**, T. Otto, and I. Symotiuk. Development of an AR-based application for assembly assistance and servicing. In *Procedia CIRP, 34th CIRP design Conference, 3-5 June 2024, Cranfield University, UK*. CIRP, 2024.

[85] A. A. Malik, T. Masood, and A. Bilberg. Virtual reality in manufacturing: immersive and collaborative artificial-reality in design of human-robot workspace. *International Journal of Computer Integrated Manufacturing*, 33(1):22–37, 2020.

[86] A. A. Malik, T. Masood, and A. Bilberg. Virtual reality in manufacturing: immersive and collaborative artificial-reality in design of human-robot workspace. *International Journal of Computer Integrated Manufacturing*, 33(1):22–37, 2020.

[87] J. Manner, S. Berg, M. Englund, B. T. Ersson, and A. Mörk. Innovative productivity improvements in forest operations: a comparative study of the assortment grapple using a machine simulator. *Journal of Forest Science*, 66(11):443–451, 2020.

[88] E. Manou, G.-C. Vosniakos, and E. Matsas. Off-line programming of an industrial robot in a virtual reality environment. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 13(2):507–519, 2019.

[89] C. Maragkos, G.-C. Vosniakos, and E. Matsas. Virtual reality assisted robot programming for human collaboration. *Procedia Manufacturing*, 38:1697–1704, 2019.

[90] J. A. Marvel, S. Bagchi, M. Zimmerman, and B. Antonishek. Towards effective interface designs for collaborative hri in manufacturing: Metrics and measures. *ACM Transactions on Human-Robot Interaction (THRI)*, 9(4):1–55, 2020.

[91] T. Masood and J. Egger. Augmented reality in support of industry 4.0—implementation challenges and success factors. *Robotics and Computer-Integrated Manufacturing*, 58:181–195, 2019.

[92] T. Masood and P. Sonntag. Industry 4.0: Adoption challenges and benefits for smes. *Computers in Industry*, 121:103261, 2020.

[93] E. Matsas and G.-C. Vosniakos. Design of a virtual reality training system for human–robot collaboration in manufacturing tasks. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 11(2):139–153, 2017.

[94] L. McAtamney and N. Corlett. Rapid upper limb assessment (rula). In *Handbook of human factors and ergonomics methods*, pages 86–96. CRC Press, 2004.

[95] Microsoft. Mixed reality toolkit (mrtk). https://github.com/microsoft/MixedRealityToolkit Unity, (Accessed on April 10 2024).

[96] P. Milgram, D. Drascic, J. J. Grodski, A. Restogi, S. Zhai, and C. Zhou. Merging real and virtual worlds. In *Proceedings of IMAGINA*, volume 95, pages 218–230, 1995.

[97] M. Moghaddam, N. C. Wilson, A. S. Modestino, K. Jona, and S. C. Marsella. Exploring augmented reality for worker assistance versus training. *Advanced Engineering Informatics*, 50:101410, 2021.

[98] M. Mondellini, S. Arlati, H. Gapeyeva, K. Lees, I. Märitz, S. L. Pizzagalli, T. Otto, M. Sacco, and A. Teder-Braschinsky. User experience during an immersive virtual reality-based cognitive task: a comparison between estonian and italian older adults with mci. *Sensors*, 22(21):8249, 2022.

[99] M. Mondellini, S. Arlati, H. Gapeyeva, K. Lees, I. Märitz, S. L. Pizzagalli, T. Otto, M. Sacco, and A. Teder-Braschinsky. User experience during an immersive virtual reality-based cognitive task: a comparison between estonian and italian older adults with mci. *Sensors*, 22(21):8249, 2022.

[100] M. Mondellini, S. Arlati, S. Pizzagalli, L. Greci, M. Sacco, and G. Ferrigno. Assessment of the usability of an immersive virtual supermarket for the cognitive rehabilitation of elderly patients: A pilot study on young adults. In *2018 IEEE 6th international conference on serious games and applications for health (SeGAH)*, pages 1–8. IEEE, 2018.

[101] M. Mondellini, S. Arlati, S. Pizzagalli, L. Greci, M. Sacco, G. Ferrigno, and M. Sacco. Assessment of the usability of an immersive virtual supermarket for the cognitive rehabilitation of elderly patients: a pilot study on young adults 2018 ieee 6th international conference on serious games and applications for health (segah), 1–8 https://doi. org/10. 1109. *SeGAH*, 84013:13, 2018.

[102] M. Mondellini, S. Pizzagalli, L. Greci, and M. Sacco. Assessment of an immersive virtual supermarket to train post-stroke patients: A pilot study on healthy people. In *Augmented Reality, Virtual Reality, and Computer Graphics: 6th International Conference, AVR 2019, Santa Maria al Bagno, Italy, June 24–27, 2019, Proceedings, Part I 6*, pages 313–329. Springer, 2019.

[103] M. M. Moniri, F. A. E. Valcarcel, D. Merkel, and D. Sonntag. Human gaze and focus-of-attention in dual reality human-robot collaboration. In *2016 12th International Conference on Intelligent Environments (IE)*, pages 238–241. IEEE, 2016.

[104] L. Monostori. Cyber-physical production systems: Roots, expectations and r&d challenges. *Procedia Cirp*, 17:9–13, 2014.

[105] P. Monteiro, H. Coelho, G. Gonçalves, M. Melo, and M. Bessa. Comparison of radial and panel menus in virtual reality. *IEEE Access*, 7:116370–116379, 2019.

[106] D. Mourtzis, V. Siatras, J. Angelopoulos, and N. Panopoulos. An augmented reality collaborative product design cloud-based platform in the context of learning factory. *Procedia Manufacturing*, 45:546–551, 2020.

[107] F. Müller, D. Jaeger, and M. Hanewinkel. Digitization in wood supply–a review on how industry 4.0 will change the forest value chain. *Computers and Electronics in Agriculture*, 162:206–218, 2019.

[108] A. Naceri, D. Mazzanti, J. Bimbo, D. Prattichizzo, D. G. Caldwell, L. S. Mattos, and N. Deshpande. Towards a virtual reality interface for remote robotic teleoperation. In *2019 19th International Conference on Advanced Robotics (ICAR)*, pages 284–289, 2019.

[109] S. Nahavandi. Industry 5.0—a human-centric solution. *Sustainability*, 11(16):4371, 2019.

[110] R. Neugebauer, S. Hippmann, M. Leis, and M. Landherr. Industrie 4.0-from the perspective of applied research. *Procedia CIRP*, 57:2–7, 2016.

[111] D. Norman. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[112] R. Nossa, M. Porro, O. Picciolini, M. Malosio, S. Pittaccio, M. Valoriani, V. Asperti, F. Clasadonte, L. Oliveto, M. Mondellini, et al. What children with neuromotor disabilities need to play with technological games. *Applied Sciences*, 11(21):9886, 2021.

[113] C. Nuzzi, S. Ghidini, R. Pagani, S. Pasinetti, G. Coffetti, and G. Sansoni. Hands-free: a robot augmented reality teleoperation system. In *2020 17th International Conference on Ubiquitous Robots (UR)*, pages 617–624, 2020.

[114] S.-K. Ong, A. Yew, N. K. Thanigaivel, and A. Y. Nee. Augmented reality-assisted robot programming system for industrial applications. *Robotics and Computer-Integrated Manufacturing*, 61:101820, 2020.

[115] L. Oppermann, F. Buchholz, and Y. Uzun. Industrial metaverse: Supporting remote maintenance with avatars and digital twins in collaborative xr environments. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–5, 2023.

[116] L. Ordile, Y. Bondarenko, S. L. Pizzagalli, V. Kuts, and T. Otto. Augmented reality interface for industrial robot control and teleoperation. *nil e H ja K*, 2021.

[117] D. Ortiz Morales, S. Westerberg, P. X. La Hera, U. Mettin, L. Freidovich, and A. S. Shiriaev. Increasing the level of automation in the forestry logging process with crane trajectory planning and control. *Journal of Field Robotics*, 31(3):343–363, 2014.

[118] M. Ostanin and A. Klimchik. Interactive Robot Programming Using Mixed Reality. *IFAC-PapersOnLine*, 51(1):50–55, 2018.

[119] M. Ostanin, S. Zaitsev, A. Sabirova, and A. Klimchik. Interactive industrial robot programming based on mixed reality and full hand tracking. *IFAC-PapersOnLine*, 55(10):2791–2796, 2022. 10th IFAC Conference on Manufacturing Modelling, Management and Control MIM 2022.

[120] J. O. Oyekan, W. Hutabarat, A. Tiwari, R. Grech, M. H. Aung, M. P. Mariani, L. López-Dávalos, T. Ricaud, S. Singh, and C. Dupuis. The effectiveness of virtual environments in developing collaborative strategies between industrial robots and humans. *Robotics and Computer-Integrated Manufacturing*, 55:41–54, 2019.

[121] T. Palonen, H. Hyyti, and A. Visala. Augmented reality in forest machine cabin. *IFAC-PapersOnLine*, 50(1):5410–5417, 2017.

[122] B. Patrão and P. Menezes. A virtual reality system for training operators. *International Journal of Online Engineering (iJOE)*, 9(S8):53–55, 2013.

[123] L. Pérez, E. Diez, R. Usamentiaga, and D. F. García. Industrial robot control and operator training using virtual reality interfaces. *Computers in Industry*, 109(1):114–120, 2019.

[124] L. Pérez, S. Rodríguez-Jiménez, N. Rodríguez, R. Usamentiaga, and D. F. García. Digital twin and virtual reality based methodology for multi-robot manufacturing cell commissioning. *Applied Sciences*, 10(10):3633, 2020.

[125] M. Peruzzini, F. Grandi, and M. Pellicciari. Exploring the potential of operator 4.0 interface and monitoring. *Computers & Industrial Engineering*, 139:105600, 2020.

[126] M. Piccinelli, A. Gagliardo, U. Castellani, and R. Muradore. Trajectory planning using mixed reality: an experimental validation. In *2021 20th International Conference on Advanced Robotics (ICAR)*, pages 982–987, 2021.

[127] S. Pizzagalli, D. Spoladore, S. Arlati, M. Sacco, and L. Greci. Hic: An interactive and ubiquitous home controller system for the smart home. In *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)*, pages 1–6. IEEE, 2018.

[128] S. Pooladvand, H. Taghaddos, A. Eslami, A. Nekouvaght Tak, and U. Hermann. Evaluating mobile crane lift operations using an interactive virtual reality system. *Journal of Construction Engineering and Management*, 147(11):04021154, 2021.

[129] M. Quandt, B. Knoke, C. Gorldt, M. Freitag, and K.-D. Thoben. General requirements for industrial augmented reality applications. *Procedia Cirp*, 72:1130–1135, 2018.

[130] A. Rassõlkin, T. Vaimann, A. Kallaste, and V. Kuts. Digital twin for propulsion drive of autonomous electric vehicle. In *2019 IEEE 60th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON)*, pages 1–4. IEEE, 2019.

[131] F. Rebelo, P. Noriega, E. Duarte, and M. Soares. Using virtual reality to assess user experience. *Human factors*, 54(6):964–982, 2012.

[132] F. Rheinberg, S. Engeser, and R. Vollmeyer. Measuring components of flow: the flow-short-scale. *Proceedings of the 1st International Positive Psychology Summit*, 2002.

[133] D. Romero, P. Bernus, O. Noran, J. Stahre, and Å. Fast-Berglund. The operator 4.0: Human cyber-physical systems & adaptive automation towards human-automation symbiosis work systems. In *Advances in Production Management Systems. Initiatives for a Sustainable World: IFIP WG 5.7 International Conference, APMS 2016, Iguassu Falls, Brazil, September 3-7, 2016, Revised Selected Papers*, pages 677–686. Springer, 2016.

[134] D. Romero, J. Stahre, and M. Taisch. The operator 4.0: Towards socially sustainable factories of the future, 2020.

[135] E. Rosen, D. Whitney, E. Phillips, G. Chien, J. Tompkin, G. Konidaris, and S. Tellex. Communicating robot arm motion intent through mixed reality head-mounted displays. In *Robotics Research*, pages 301–316. Springer, 2020.

[136] P. Rückert, L. Wohlfromm, and K. Tracht. Implementation of virtual reality systems for simulation of human-robot collaboration. *Procedia Manufacturing*, 19(1):164–170, 2018.

[137] T. Schubert, F. Friedmann, and H. Regenbrecht. Igroup presence questionnaire. *Teleoperators Virtual Environ.*, 41:115–124, 2001.

[138] C. X. E. (Shamaine), Y. Qiao, V. Kuts, J. Henry, K. McNevin, and N. Murray. Teleoperation of the industrial robot: Augmented reality application. In *Proceedings of the 13th ACM Multimedia Systems Conference*, MMSys '22, page 299–303, New York, NY, USA, 2022. Association for Computing Machinery.

[139] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson. Virtual reality induced symptoms and effects (vrise): Comparison of head mounted display (hmd), desktop and projection display systems. *Displays*, 29(2):58–69, 2008.

[140] S. Shirowzhan, W. Tan, and S. M. Sepasgozar. Digital twin and cybergis for improving connectivity and measuring the impact of infrastructure construction planning in smart cities. *ISPRS International Journal of Geo-Information*, 9(4):240, 2020.

[141] M. Simonetto, M. Peron, G. Fragapane, and F. Sgarbossa. Digital assembly assistance system in industry 4.0 era: a case study with projected augmented reality. In *Advanced Manufacturing and Automation X 10*, pages 644–651. Springer, 2021.

[142] J. E. Solanes, A. Muñoz, L. Gracia, A. Martí, V. Girbés-Juan, and J. Tornero. Teleoperation of industrial robot manipulators based on augmented reality. *The International Journal of Advanced Manufacturing Technology*, 111:1077–1097, 2020.

[143] M. Speicher, A. M. Feit, P. Ziegler, and A. Krüger. Selection-based text entry in virtual reality. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–13, 2018.

[144] F. Stramandinoli, K. G. Lore, J. R. Peters, P. C. O'Neill, B. M. Nair, R. Varma, J. C. Ryde, J. T. Miller, and K. K. Reddy. Robot learning from human demonstration in virtual reality. In *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI)*, 2018.

[145] I. E. Sutherland et al. The ultimate display. In *Proceedings of the IFIP Congress*, volume 2, pages 506–508. New York, 1965.

[146] S. K. Tadeja, Y. Lu, P. Seshadri, and P. O. Kristensson. Digital twin assessments in virtual reality: An explorational study with aeroengines. In *2020 IEEE Aerospace Conference*, pages 1–13, Big Sky, Montana, USA, 2020. IEEE, IEEE.

[147] W. Terkaj, E. Pessot, V. Kuts, Y. Bondarenko, S. L. Pizzagalli, and K. Kleine. A framework for the design and use of virtual labs in digital engineering education. In *AIP Conference Proceedings*, volume 2989-1. AIP Publishing, 2024.

[148] W. Terkaj, T. Tolio, and M. Urgo. A virtual factory approach for in situ simulation to support production and maintenance planning. *CIRP Annals*, 64(1):451–454, 2015.

[149] **S.L. Pizzagalli**, Y. Bondarenko, B. C. Baykara, A. Niidas, V. Kuts, M. Kerm, and T. Otto. Forestry crane immersive user interface for control and teleoperation. In *ASME International Mechanical Engineering Congress and Exposition*, volume 86649, page V02BT02A021. American Society of Mechanical Engineers, 2022.

[150] **S.L. Pizzagalli**, Y. Bondarenko, V. Kuts, E. O'Connell, N. Murray, and T. Otto. Ros-based augmented and virtual reality path planning interface for industrial robotic arms: a preliminary assessment. In *ACM, the 6th International Conference on Robot Systems and Applications (ICRSA), 2023*. ACM, 2024.

[151] **S.L. Pizzagalli**, V. Kuts, Y. Bondarenko, and T. Otto. Evaluation of virtual reality interface interaction methods for digital twin industrial robot programming and control, a pilot study. In *ASME International Mechanical Engineering Congress and Exposition*, volume 85567, page V02BT02A005. American Society of Mechanical Engineers, 2021.

[152] **S.L. Pizzagalli**, V. Kuts, and T. Otto. User-centred design in industrial collaborative automated systems. *Proceedings of the Estonian Academy of Sciences*, 70(4):436–443, 2021.

[153] T. Togias, C. Gkournelos, P. Angelakis, G. Michalos, and S. Makris. Virtual reality environment for industrial robot control and path design. *Procedia CIRP*, 100:133–138, 2021. 31st CIRP Design Conference 2021 (CIRP Design 2021).

[154] S. Vaidya, P. Ambad, and S. Bhosle. Industry 4.0–a glimpse. *Procedia manufacturing*, 20:233–238, 2018.

[155] H. Van der Valk, H. Haße, F. Möller, M. Arbter, J.-L. Henning, and B. Otto. A taxonomy of digital twins. In *AMCIS*, 2020.

[156] D. L. Vickers. *Sorcerer's apprentice: head-mounted display and wand*. PhD thesis, Dept. of Electrical Engineering, University of Utah., 1972.

[157] A. Vidal-Balea, O. Blanco-Novoa, P. Fraga-Lamas, M. Vilar-Montesinos, and T. M. Fernández-Caramés. Creating collaborative augmented reality experiences for industry 4.0 training and assistance applications: Performance evaluation in the shipyard of the future. *Applied Sciences*, 10(24):9073, 2020.

[158] V. Villani, F. Pini, F. Leali, and C. Secchi. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics*, 55:248–266, 2018.

[159] V. Villani, F. Pini, F. Leali, C. Secchi, and F. Cesare. Survey on Human-Robot Interaction for Robot Programming in Industrial Applications. *IFAC-PapersOnLine*, 51(1):66–71, 2018.

[160] G.-C. Vosniakos, L. Ouillon, and E. Matsas. Exploration of two safety strategies in human-robot collaborative manufacturing using virtual reality. *Procedia manufacturing*, 38:524–531, 2019.

[161] B. Wang, P. Zheng, Y. Yin, A. Shih, and L. Wang. Toward human-centric smart manufacturing: A human-cyber-physical systems (hcps) perspective. *Journal of Manufacturing Systems*, 63:471–490, 2022.

[162] Q. Wang, W. Jiao, P. Wang, and Y. Zhang. Digital twin for human-robot interactive welding and welder behavior analysis. *IEEE/CAA Journal of Automatica Sinica*, 8(2):334–343, 2020.

[163] Y. Wang, G. Zhai, S. Chen, X. Min, Z. Gao, and X. Song. Assessment of eye fatigue caused by head-mounted displays using eye-tracking. *Biomedical engineering online*, 18:1–19, 2019.

[164] S. Westerberg and A. Shiriaev. Virtual environment-based teleoperation of forestry machines: Designing future interaction methods. *Journal of Human-Robot Interaction*, 2(3):84–110, 2013.

[165] D. Whitney, E. Rosen, E. Phillips, G. Konidaris, and S. Tellex. Comparing robot grasping teleoperation across desktop and virtual reality with ros reality. In *Robotics Research: The 18th International Symposium ISRR*, pages 335–350. Springer, 2019.

[166] G. A. Yashin, D. Trinitatova, R. T. Agishev, R. Ibrahimov, and D. Tsetserukou. Aerovr: Virtual reality-based teleoperation with tactile feedback for aerial manipulation. In *2019 19th International Conference on Advanced Robotics (ICAR)*, pages 767–772. IEEE, 2019.

[167] H. Yun and M. B. Jun. Immersive and interactive cyber-physical system (i2cps) and virtual reality interface for human involved robotic manufacturing. *Journal of Manufacturing Systems*, 62:234–248, 2022.

Acknowledgements

A special thank to the supervisors, colleagues and friends, and everyone caring, aiding this journey towards accomplishment with suggestions, recommendations, opinions, and precious discussions.

The research was conducted using the Smart Industry Centre (SmartIC) core facility funded by the Estonian Research Council grant TT2 and supported by the "Secure 5G-Enabled Twin Transition for Europe's TIMBER Industry Sector" project funded by European Union's Horizon Europe Research program (5G-TIMBER).

Thanks to my kids, as they never stopped loving and supporting me.

Abstract

Extended Reality for I5.0: Towards Human Centricity in Human-Robot Interaction

Contemporary industrial paradigms and technological developments are reducing the gaps, and physical spaces, between operators and machines, requiring new collaborative methods and interfaces for Human Robot Interaction (HRI) together with the inclusion of humans in the loop for twinned and cyber-physical systems. The development of the Industry5.0 (I5.0) paradigm leads to a re-evaluation of the role of workers towards, supposedly, a focus on safety, well-being and user centric approach.

This scenario challenges the prevalent development methods and tools, requiring an integrated and multidisciplinary viewpoint to allow creativity and automation to converge in the expected flexible and resilient production paradigm. The interaction and integration of technologies, interfaces, design and assessment procedures focusing on operators rather than data exploitation and automation only, seem to be the only way to move forward. This is rarely the case both in research and industrial implementation, where, despite the widespread integration of Digital Twins (DT) and Extended Reality (XR) in different use cases, the two technologies still struggle when it comes to full integration, especially in Small Medium Enterprises (SMEs), and endorsement of effective development and assessment methods.

In this context a User Centered Design (UCD) approach is adopted to lead the proposed consolidation of Digital Twins, Human in the Loop and Extended Reality interfaces components towards a more efficient, safe and reliable Human Robot Interaction (HRI). The validation of DT based XR interfaces for HRI is arranged along the UCD loop and based on the unique Human Computer Interaction happening within and without the digital environments.

The development and assessment of such systems took place in laboratory based user studies, comparing DT XR interfaces against native controls on robotic arm manipulation and teleoperation for pick-and-place and path planning tasks. These experiments encompassed evaluations of performance, efficiency, effectiveness, satisfaction, user experience, task load and system precision with the adoption of quantitative, qualitative and subjective metrics revolving around the end users.

Results confirm that XR interfaces are at least as reliable and precise as the physical controls counterparts with trends of lower task load, a better user experience, higher usability and efficiency especially for fully immersive headset based scenarios. These are generally resulting in higher reliability and usability scores compared to augmented reality headsets. The adopted design and assessment approach has been successfully validated in two real world use cases in local SMEs active in the forestry, and wood processing field. Different XR applications aimed at support of operators in machine control, maintenance and product assembly tasks have been developed and partially evaluated with the end users along the line of the laboratory experiences.

In conclusion, the presented studies, along with a proposed research plan following the UCD loop, aim to facilitate the advancements in the context of Operator4.0 and human cyber-physical systems design and assessment, addressing the current demand for flexible, resilient, decentralized, and user-centric solutions in industry, thus potentially accelerating the adoption of DT XR technologies in SMEs by offering structured procedures and tools while contributing to the refinement of standards and metrics for XR-based Human-Robot Interaction.

Kokkuvõte
Laiendatud reaalsus Tööstus 5.0 jaoks: inimesekeskse lähenemiseni inimese-roboti suhtluses

Nüüdisaegsed tööstusparadigmad ja tehnoloogiline areng vähendavad lõhesid ja füüsilist ruumi operaatorite ja masinate vahel, nõudes uusi koostöömeetodeid ja liideseid inimese ja roboti vaheliseks suhtluseks (HRI), kaasates inimesi kaasamisega kaksik- ja küberfüüsilistesse süsteemidesse. Tööstus 5.0 (I5.0) paradigma areng viib töötajate rolli ümberhindamiseni, keskendudes eeldatavalt ohutusele, heaolule ja kasutajakesksele lähenemisele.

See stsenaarium seab väljakutse valdavatele arendusmeetoditele ja -tööriistadele, nõudes integreeritud ja multidistsiplinaarset vaatenurka, mis võimaldaks loovusel ja automatiseerimisel konvergeeruda oodatavas paindlikus ja vastupidavas tootmisparadigmas. Tehnoloogiate, liideste, disaini ja hindamisprotseduuride koostoime ja integreerimine, keskendudes pigem operaatoritele kui ainult andmete ärakasutamisele ja automatiseerimisele, tundub olevat ainus viis edasi liikuda. See on harva nii uurimistöös kui ka tööstuslikus rakenduses, kus hoolimata digitaalsete kaksikute (DT) ja laiendatud reaalsuse (XR) ulatuslikust integreerimisest erinevates kasutusjuhtumites, on mõlemal tehnoloogial raskusi täieliku integreerimisega, eriti väikestes ja keskmise suurusega ettevõtetes (VKE-des), ning tõhusate arendus- ja hindamismeetodite toetamisega.

Selles kontekstis võetakse kasutusele kasutajakeskne disain (UCD), et juhtida väljapakutd digitaalsete kaksikute, inimese kaasamise ja laiendatud reaalsuse liideste komponentide planeeritud konsolideerimist tõhusama, turvalisema ja usaldusväärsema inimese-roboti suhtluse suunas. DT-põhiste XR-liideste HRI valideerimine toimub kasutajakeskse disaini tsükli raames ning põhineb unikaalsel inimese-arvuti suhtlusel nii digitaalsetes keskkondades kui ka nende väljaspool.

Selliste süsteemide arendus ja hindamine toimusid laboripõhistes kasutajauuringutes, võrreldes DT XR-liideste jõudlust traditsiooniliste juhtimispultidega robotkäe manipuleerimisel ja kaugjuhtimisel nii valiku- ja asetamistööde kui ka teekonna planeerimise ülesannete puhul. Need eksperimendid hõlmasid tulemuslikkuse, efektiivsuse, rahulolu, kasutajakogemuse, ülesandekoormuse ja süsteemi täpsuse hindamist, võttes kasutusele kvantitatiivsed, kvalitatiivsed ja subjektiivsed meetrikaid, mis keskenduvad lõppkasutajatele.

Tulemused kinnitavad, et XR-liidesed on vähemalt sama usaldusväärsed ja täpsed kui füüsilised juhtseadised, näidates madalamat ülesandekoormust, paremat kasutajakogemust, suuremat kasutatavust ja efektiivsust eriti täielikult immersiivsete peakomplektidega stsenaariumides. Need tulemused annavad üldiselt suurema usaldusväärsuse ja kasutatavuse hinded võrreldes laiendatud reaalsuse peakomplektidega. Kasutatud disaini- ja hindamislähenemist on edukalt valideeritud kahe reaalse maailma kasutusjuhtumi puhul kohalikes VKE-des, mis tegutsevad metsanduse ja puidutöötlemise valdkonnas. Erinevaid XR rakendusi, mis on suunatud operaatorite toetamisele masinate juhtimises, hoolduses ja toote koostamise ülesannetes, on arendatud ja osaliselt hinnatud koos lõppkasutajatega vastavalt laborikogemustele.

Kokkuvõttes on esitatud uuringud koos ettepanekuga järgida kasutajakeskse disaini tsüklit, eesmärgiga soodustada edusamme Operaator 4.0 ja inimese küberfüüsiliste süsteemide disaini ja hindamise kontekstis, vastates tööstuse praegusele nõudlusele paindlike, vastupidavate, detsentraliseeritud ja kasutajakesksete lahenduste järele, kiirendades potentsiaalselt DT XR-tehnoloogiate kasutuselevõttu VKE-des, pakkudes struktureeritud protseduure ja tööriistu ning aidates kaasa standardite ja meetrikate täiustamisele XR-põhise inimese-roboti suhtluse jaoks.

I publication

I

S.L. Pizzagalli, V. Kuts, and T. Otto. User-centred design in industrial collaborative automated systems. *Proceedings of the Estonian Academy of Sciences*, 70(4):436–443, 2021

Proceedings of the Estonian Academy of Sciences,
2021, **70**, 4, 436–443
https://doi.org/10.3176/proc.2021.4.10
Available online at www.eap.ee/proceedings

MANUFACTURING
ENGINEERING

User-centred design in industrial collaborative automated systems

Simone Luca Pizzagalli*, Vladimir Kuts and Tauno Otto

Department of Mechanical and Industrial Engineering, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia

Received 15 June 2021, accepted 16 July 2021, available online 3 November 2021

Abstract. Autonomous systems and collaborative robotics are part of the pillar technologies of the Industry 4.0 (I4.0) paradigm. These include advanced simulations, Digital Twins (DTs) and novel Human Machine Interfaces (HMIs). The increasing development of these technologies together with the higher requirements for customized production processes demands a closer collaboration between operators and automated systems. This leads to a redefinition of how human operators manage and interact with machines and how they are supported in this by adaptable interfaces, simulations and real-time data collection and analysis. New Human-Robot Collaboration (HRC) paradigms are paramount in a scenario where the boundaries between human and machine performed tasks are flexible and increasingly dematerialized. The redefinition of standards, design methods, programming interfaces and assessment techniques is central to facilitate these technological and production changes. The augmentation of human capabilities in the workplace insists on a definition of a framework of requirements that would integrate human, organizational and production needs in the same scenario and workflow. This research proposes a User-Centred Design (UCD) approach which is crucial in addressing the open challenges of HRC systems. Our work regards the DT as well as Augmented and Virtual Reality (AR/VR) technologies as central in this process by considering them key tools for the design, control, and assessment of modern collaborative industrial scenarios.

Key words: human-robot collaboration, user-centred design, Operator 4.0, digital twins, virtual reality.

1. INTRODUCTION

The Industry 4.0 (I4.0) revolution has systematized and clarified the future scenario for production processes towards the development of smart efficient and automated systems, including in its leading technologies big data and analytics, autonomous robots, simulation, system integration, industrial internet of things, cyber-physical systems, the cloud, additive manufacturing, and Augmented and Virtual Reality (AR/VR) [1]. Human-Robot Collaboration (HRC) is a crucial aspect in this scenario. Recent advancements in industrial automated systems and the need of production line flexibility, adaptability to market demand and customization are leading to a closer collaboration routine between human operators and robots in several industrial fields. The boundaries between the operator and machines are becoming less evident and tangible while the technologies involved in the modern production processes have a direct impact on the human workers, as they support and perfect human skills by sensing and perceptive technologies throughout the production lifecycle [2]. The newly born teaming paradigms are in need of natural, adaptive and intuitive Human Machine Interfaces (HMIs) that could support safety in HRC but also promote wellbeing and technology acceptance from the human operator perspective [3]. Recent studies show how Digital Twins (DTs) AR/VR interfaces can be valuable tools in the evaluation of safety [4] programming methods [5] and commissioning of HRC systems, while only a few studies attempt the combination of DT AR/VR technologies in a User-Centred Design (UCD) approach focusing on specific use cases. As much

* Corresponding author, simone.pizzagalli@taltech.ee

as industries adopting and implementing I4.0 technologies in their production lines are aware of the organization and human-related impacts and risks [6], no clear design and assessment methods are yet available or sufficient to provide a robust framework that would include human factors in design and decision-making strategies related to HRC. This research aims to address the HRC challenges by integrating the industrial collaborative robotic technologies into a user-centred design and evaluation framework. The main research question is whether DT AR/VR interfaces could be used as central tools for the development of modern user-centred robot collaboration systems, improving the efficiency and safety of interaction between operators and machines in the manufacturing field.

2. HRC TRENDS IN I4.0

Modern HRC systems combine the advantages of automation task repeatability and workload with specialized workers' soft skills and decision making [7]. Nevertheless, there are some requirements and open challenges in the development of efficient collaborative technologies including the design and implementation of intuitive interfaces. The latter would include different input modes, outputs and feedback, and robot centric issues such as safety, diversified programming approaches, design and control methods [8]. The future massive exploitation of collaborative robots in the manufacturing field will lead to many production processes and organizational adaptations within the industries and a series of human-centred concerns [9]. The impact of these technologies extends from the industrial environment to the operators with many different implications in each step of the production system. A large number of challenges are open to evaluation, assessment and integrated design solutions that would ease or accommodate operator monitoring technologies, stress level and workload assessment, technology acceptance, efficient reprogramming, ethics, privacy and handling of the data collected on the shop floor [10]. In this sense, the role of the operator itself and the type of interaction with the automated technology will be radically shifting from cooperation to an augmented and symbiotic paradigm. Romero et al. [11] describe this new worker as the Operator 4.0 (O4.0), enhanced in its physical, sensorial and cognitive capabilities by the same technologies and systems involved in the fourth industrial revolution. This scenario requires the agents involved in this new production paradigm to be intelligent, perceptive and purposeful, aware of the context, autonomous, able to act, reflective, adaptable, learning, and conversational. Many of these characteristics are part of the human skills and competencies but seemingly missing in any industrial robotic system. Technologies that are already employed in advanced manufacturing support this evolution to a symbiotic enhancement of human operators' capabilities, being at the same time the main causes of ethical and human-related issues. IoT, advanced sensors, artificial intelligence, and data analysis promote smarter safety systems and learning approaches in robotic collaboration and programming by grounding their advanced capabilities precisely on human and environment-based data and behaviour analysis. Human operators are becoming a central matter of concern and the main driver of this transition. New methods are necessary to efficiently integrate the I4.0 technologies with the new O4.0 paradigm in a sustainable way. What is missing is a systematic approach that would allow the evaluation of all aspects involved in the HRC process and make DT AR/VR interfaces the main tool bridging humans and robots in a UCD approach.

2.1. Collaboration levels, safety, and programming in HRC

Levels of collaboration between the operator and automated systems, safety measures and programming methods, which are relying on advanced User Interface (UI), are among the most important aspects of HRC systems. Collaborative operations between the human partner and the robot can have different characterizations. In the study by Helms et al. [12] four types of levels are described: independent or parallel work, synchronized work, simultaneous work and assisted or collaborative scenarios. The latter is crucial in our proposed architecture and involves sharing workpiece, tools and workspace while performing the task at the same time. The relationship between safety, coexistence and collaboration in Human-Robot Interaction (HRI) is defined by De Luca et al. [13] as nested levels of the same framework describing interactions between humans and robots. While safety is considered the basic feature of industrial automated systems, coexistence and collaboration describe levels of interaction which involve sharing the same space and direct physical contact between the operator and the machine. Safety requirements for industrial collaborative automated systems are described in [14] and include four main different methods: Safety-Rated Monitored Stop, Hand Guiding, Speed and Separation Monitoring, and Power and Force Limiting. Each of these methods suggest a varied type of involvement and level of proximity which have various implications in task performance and programming operations. Different interaction methods must be supported by suitable context awareness approaches allowing automated systems to behave intelligently within the shop floor. These are mediated by artificial intelligence, industrial internet of things and advanced sensors which have different impact on the user, depending on the specific use case scenario and interaction modality. Programming of modern collaborative robots involves a variety of dedicated interfaces and methods. The most common techniques include: offline programming, where the operations are simulated on a computer and eventually exported to the robot system; online programming which involves traditional teach pendant interfaces; walk-through programming, a common method in the last generation of Cobots, in which the operator directly manipulates the robot while paths and trajectories are saved by the systems; Programming by Demonstration (PbD), where the robotic systems learn the operations from actions performed by a human teacher. As much as not all robotic systems are suitable to be directly manipulated, the need for precise and reliable sensorized suites or motion tracking techniques seems to add complexity and costs to the adoption of efficient PbD methods. Moreover, world state information beyond the teacher-operator or the robot itself is not always properly mapped and described by the system [15]. AR/VR technologies offer a solution to several issues involved in programming procedures. The human operator is a predominant agent in all programming techniques by either manipulating the robot, demonstrating and teaching trajectories and operations, or as the end-user of complex programming software and UIs. Virtual reality interfaces offer a repeatable, scalable, controlled and safe test ground for innovative programming methods such as kinesthetic teaching [16] and user monitoring, and it can be applicable to older generations of robots.

2.2. DT and AR/VR interfaces in industrial automated system scenarios

DTs are advanced representations of the real-world state and product or system components allowing for a real-time synchronized loop transmission of data between the digital and real entities. Beyond the first NASA applications and experimental works in this field and the well-known definition provided by Grieves [17], many other characterizations have been given to this technology [18]. The popularity of DT solutions both in the industrial and research field is constantly increasing with examples spanning from manufacturing applications, electric engine optimization [19] to building construction [20], aviation, and healthcare [21]. The flexibility of simulation-driven systems allows for applications supporting maintenance and production planning [22] to real-time data-driven models for product monitoring [23]. A few examples of a user-inclusive DT perspective in industrial systems can be found in [24–26]. Nevertheless, these approaches seems to be limited, both in terms of inclusion of the human factors in the loop, and in the design of accessible and intuitive interfaces aimed at data visualization and interactions with the digital counterpart [18]. AR/VR technologies can merge, on different levels and with different degrees of blending, the virtual and the real world [27], facilitating data visualization and creating a means of interaction between simulated systems and their real counterpart. A major advantage of AR/VR HMI for HRC is the inclusion of the end-user in the robotic cell's User Interface. By being an active agent in the automated system control UI, which in some cases coincides with the DT of the systems and factory floor environment itself, the operator becomes part of the twinning loop. This allows for direct data collection, task repeatability, user and environmental monitoring. Many recent studies have focused on experimental approaches to DT AR/VR interfaces with applications aimed at programming [28], control [29], design of collaborative industrial cells [30], and assessment of safety in HRC systems [31]. Only a limited number of examples address the potential of these technologies in becoming UCD and evaluation tools for advanced industrial systems and workstations [6,32]. Other experimental works, [4,33,34], attempt to adopt several user and system evaluation metrics for the assessment of immersive DT UI and interactions. These examples are limited to specific use cases and are not providing sufficiently comprehensive study protocols for HRI both from the technological system's and user's point of view. The shifting roles of automated technologies and human workforce in modern industry need a design and evaluation framework that can allow testing and validating interface efficiency for robot control, real-time system monitoring and synchronization, and impact on the user.

3. UCD FOR HRC

A different design approach is indispensable to efficiently integrate the new paradigms of I4.0 and O4.0 requirements into HRC. This framework should be able to integrate into the same design scope the above-mentioned challenges of collaborations methods, task organization, safety, programming, accessible UIs, production process organization and technology impact on the operator and factory organization. We propose that this integration should happen along the UCD iterative design process as defined in [35]. This design method places the end users at the centre and all along the design process involves different stakeholders and multidisciplinary resources in planning, testing, commissioning, and assessment of the system. Open multidisciplinary approaches can be found in design of complex intelligent systems such as autonomous vehicles [36]. Requirements should be drawn by analysing the target user group, the specific context of use, including organizational needs and objectives. Personal information, as well as demographics, knowledge of technical solutions, health-related issues, or body functionality, are at the base of the characterization of generic user models involved in the preliminary design process. Ethnographic research methods, including questionnaires, interviews, surveys, direct observations, can help in clarifying the needs and behaviours of the involved stakeholders at different levels of the context of use and design phases. Figure 1 shows how the new design approach allocates safety, collaboration modes, interfaces, task oriented and managerial requirements along the proposed iterative design and assessment process, matching the requirements and technologies which are typically involved in HRC with the UCD loop and specifications.

Our research proposal puts DT AR/VR interfaces at the centre of this design method constituting the O4.0 layer in the HRC structure and favouring interactions with the robotics systems. Figure 2 demonstrates the UCD

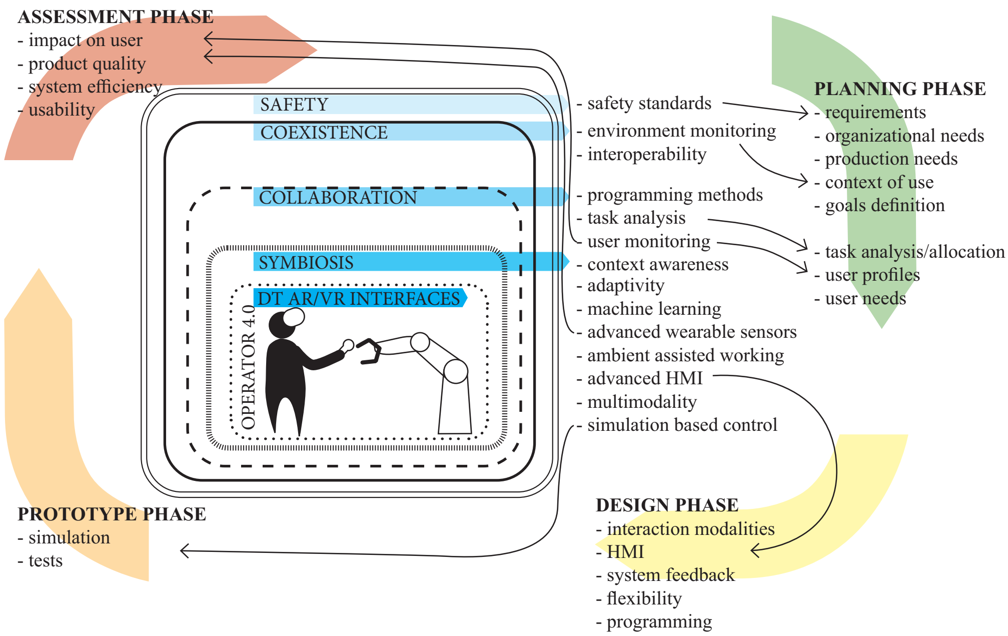

Fig. 1. Synergies between modern HRC systems and a user-centred iterative design process.

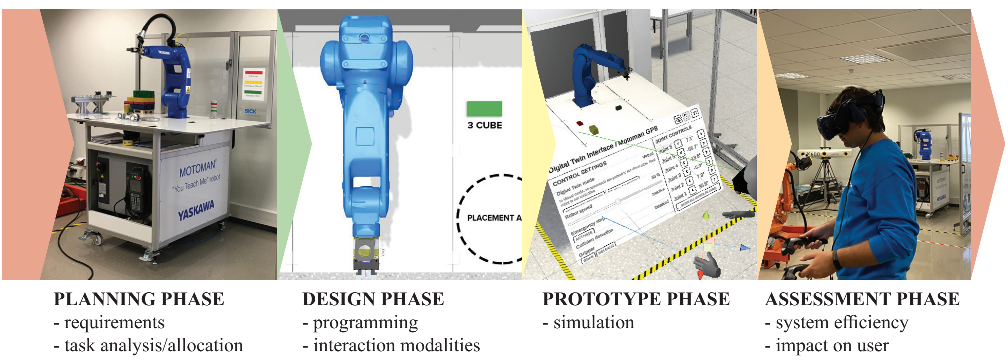

Fig. 2. Use case scenario of UCD approach in HRC making use of DT AR/VR interfaces.

approach applied to the use case developed and discussed in [37]. This industrial robot use case shows how DT AR/VR technologies are able to include the operator in the HRC system interfaces.

Assessment of users' environmental and organizational aspects can be achieved by adopting real-time simulations and interaction technologies. These are based on industrial IoT, artificial intelligence, advanced sensors and data analysis having human, environment data and behaviour analysis as the main target, and smart, safe and easy-to-manage collaborative robotic systems as the main goal.

4. USER-CENTRED DT AR/VR HRC SYSTEM ARCHITECTURE

DT AR/VR technologies demonstrate high adaptability and scalability towards different features, programming methods and interfaces of industrial automated systems. Immersive simulations include the user in the DT loop and integrate environmental state information, allowing for safe reproducible and controlled experimental test grounds for the configuration and design of cyber-physical systems. For this reason, VR technologies could be adopted in existing ethnographic research methods [38]. The fast-paced advancements in hardware technologies favour the integration of motion tracking, eye-tracking, environmental or physiological sensors with VR head-mounted displays and controllers. This enables many different interaction modalities, gestures, vocal or motion control with the digitalized counterpart of real machinery as well as the possibility of adopting advanced UI adaptivity and adaptability methods. Figure 3 shows the software integration model for DT AR/VR based collaborative automated system, structured along a UCD approach.

This model contributes to a highly immersive VR/AR training and control system. It provides an agent with higher-level involvement than customary training or control, and thus being more efficient in the HRC process. As already mentioned, there is a need for new metrics to assess and validate DT AR/VR interfaces for manufacturing systems that would respond to advancements in I4.0 technologies and new production requirements. The central point of the architecture and the proposed approach is that the metrics and methods for advanced collaborative systems should be developed in DT immersive interfaces first. VR interactive systems constitute the link between interaction and simulation by being at the same time the technology that allows humans to interact with the simulation and the HRC interface itself. In this scenario a further clarification is necessary on how one interacts within the virtual world, based on Human Computer Interaction (HCI) methodologies and a categorization of interaction metaphors and illusions allowed by the simulated world. It should also be clarified what the consequences of the interactions are on the DT of the robotic system in the simulation, based on DT technology enablers, requirements and performance; and how one interacts with the computer system and technology allowing for the simulation and interaction in the DT. The latter is also based on HCI standards but has proved to have both unprecedented built-in multimodal

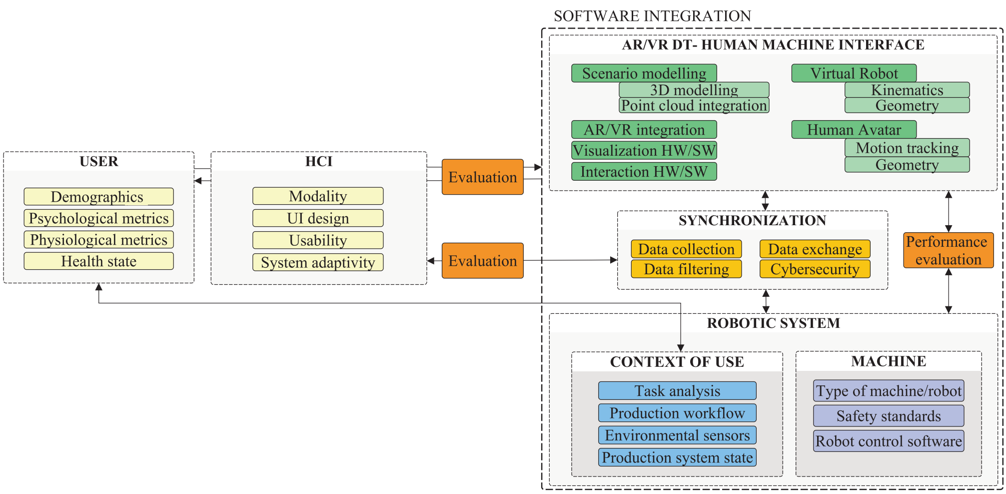

Fig. 3. Software integration architecture for UCD based HRC systems adopting DT AR/VR interfaces.

potential and impact on the user. The proposed architecture responds to these requirements by efficiently allocating evaluation of DT system performance, HMI efficacy and usability, impact on the user and HRI assessment.

5. CONCLUSIONS

Holistic design solutions are necessary to respond to the impelling changes in industrial automated systems and the newly gained central position of the operators in collaborative scenarios. In this context human stakeholders, context of use, and organization of the production systems need to be reconsidered and optimized to match the new industrial requirements. UCD demonstrates how these elements can be organized in a design process that would iteratively plan, design and assess the newly adopted solutions while taking into account all the above-mentioned critical elements. The central technologies in this process are DT AR/VR interfaces which have already demonstrated their scalability and efficiency in robotic system control and interaction, to support the evolution of O4.0 capabilities on the shop floor. As already mentioned, immersive visualization and interaction technologies include the user in the simulation and system interface supporting the evaluation of metrics and methods for HMI in HRC scenarios from UCD perspective. This new design and assessment solution would facilitate both the evaluation of psychological and physiological well-being ergonomics factors in HRI and the efficiency of advanced production processes and industrial cell deployment. Future research will give an in-depth overview of the state of the art of the DT AR/VR applications in the industrial field by comparing the type of use case scenarios, interaction methods, and assessment metrics. Based on the existing use case presented in Fig. 2, future experimental studies will try to assess the interaction of the human operator with each layer of the proposed architecture including the interaction with the robot, with the AR/VR technologies and hardware, and the DT simulation of the industrial system. Future user-centred studies will find shared evaluation metrics between the different interaction levels, the design solutions of which would allow best performance and low impact on the user and on how AR/VR are best integrated with the DT loop in HRC scenarios. Interrelations between the elements of the proposed architecture need to be evaluated in light of the defined assessment parameters. Customization and visualization of these connections might also be taken into consideration for a better fit into different use case scenarios. From the technical point of view cyber-security for user and environmental data collection and handling will have to be addressed. Connectivity and synchronization methods between the physical and virtual systems will be further explored and extended to include virtual distributed infrastructure for control and interaction with different machinery that will offer an experimental test ground and modular virtual interface for both system integrators and researchers.

ACKNOWLEDGEMENTS

This work was supported by the Estonian Research Council grant TT2 “Smart Industry Centre (SmartIC)”. The publication costs of this article were covered by the Estonian Academy of Sciences and Tallinn University of Technology.

REFERENCES

1. Vaidya, S., Ambad, P. and Bhosle, S. Industry 4.0 − a glimpse. *Procedia Manuf.*, 2018, **20**, 233−238. https://doi.org/10.1016/j.promfg.2018.02.034
2. Romero, D., Bernus, P., Noran, O., Stahre, J. and Fast-Berglund, Å. The Operator 4.0: human cyber-physical systems & adaptive automation towards human-automation symbiosis work systems. *IFIP Adv. Inf. Commun. Technol.*, 2016, **488**, 677−686. https://doi.org/10.1007/978-3-319-51133-7_80
3. Berg, J. and Lu, S. Review of interfaces for industrial human-robot interaction. *Curr. Rob. Rep.*, 2020, **1**, 27−34. https://doi.org/10.1007/s43154-020-00005-6
4. Oyekan, J. O., Hutabarat, W., Tiwari, A., Grech, R., Aung, M. H., Mariani, M. P. et al. The effectiveness of virtual environments in developing collaborative strategies between industrial robots and humans. *Rob. Comput. Integr. Manuf.*, 2019, **55**(A), 41−54. https://doi.org/10.1016/j.rcim.2018.07.006
5. Burghardt, A., Szybicki, D., Gierlak, P., Kurc, K., Pietruś, P. and Cygan, R. Programming of industrial robots using virtual reality and digital twins. *Appl. Sci*., 2020, **10**(2), 486. https://doi.org/10.3390/app10020486
6. Longo, F., Padovano, A. and Umbrello, S. Value-oriented and ethical technology engineering in Industry 5.0: A human-centric perspective for the design of the factory of the future. *Appl. Sci.*, 2020, **10**(12), 4182. https://doi.org/10.3390/APP10124182
7. Nahavandi, S. Industry 5.0 − a human-centric solution. *Sustainability*, 2019, **11**(16), 4371. https://doi.org/10.3390/su11164371
8. Villani, V., Pini, F., Leali, F. and Secchi, C. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics*, 2018, **55**, 248−266. https://doi.org/10.1016/j.mechatronics.2018.02.009
9. Demir, K. A., Döven, G. and Sezen, B. Industry 5.0 and human-robot co-working. *Procedia Comput. Sci.*, 2019, **158,** 688−695. https://doi.org/10.1016/j.procs.2019.09.104
10. Bagdasarov, Z., Martin, A. A. and Buckley, M. R. Working with robots: Organizational considerations. *Organ. Dyn.*,

2020, **49**(2), 100679. https://doi.org/10.1016/j.orgdyn.2018.09.002
11. Romero, D., Stahre, J. and Taisch, M. The Operator 4.0: Towards socially sustainable factories of the future. *Comput. Ind. Eng.*, 2020, **139**, 106128. https://doi.org/10.1016/j.cie.2019.106128
12. Helms, E., Schraft, R. D. and Hägele, M. Rob@work: Robot assistant in industrial environments. In *Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication, Berlin, Germany, September 27, 2002*. IEEE, 399−404. https://doi.org/10.1109/ROMAN.2002.1045655
13. De Luca, A. and Flacco, F. Integrated control for pHRI: Collision avoidance, detection, reaction and collaboration. In *Proceedings of the 2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob), Rome, Italy, June 24−27, 2012*. IEEE, 288–295. https://doi.org/10.1109/BioRob.2012.6290917
14. ISO/TS 15066:2016. Robots and robotic devices – Collaborative robots. International Organization for Standardization, Geneva, 2016. https://www.iso.org/standard/62996.html
15. Argall, B. D., Chernova, S., Veloso, M. and Browning, B. A survey of robot learning from demonstration. *Rob. Auton. Syst.*, 2009, **57**(5), 469−483. https://doi.org/10.1016/j.robot.2008.10.024
16. DelPreto, J., Lipton, J. I., Sanneman, L., Fay, A. J., Fourie, C., Choi, C. and Rus, D. Helping robots learn: a human-robot master-apprentice model using demonstrations via virtual reality teleoperation. In *Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, May 31−August 31, 2020*. IEEE, 10226–10233. https://doi.org/10.1109/ICRA40945.2020.9196754
17. Grieves, M. Digital Twin: manufacturing excellence through virtual factory replication. White Paper, 2014. https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_through_Virtual_Factory_Replication
18. Barricelli, B. R., Casiraghi, E. and Fogli, D. A survey on Digital Twin: definitions, characteristics, applications, and design implications. *IEEE Access*, 2019, **7**, 167653−167671. https://doi.org/10.1109/ACCESS.2019.2953499
19. Rassõlkin, A., Vaimann, T., Kallaste, A. and Kuts, V. Digital twin for propulsion drive of autonomous electric vehicle. In *Proceedings of the 2019 IEEE 60th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), Riga, Latvia, October 7−9, 2019*. IEEE, 2020, 1−4. https://doi.org/10.1109/RTUCON48111.2019.8982326
20. Khajavi, S. H., Motlagh, N. H., Jaribion, A., Werner, L. C. and Holmström, J. Digital Twin: vision, benefits, boundaries, and creation for buildings. *IEEE Access*, 2019, **7**, 147406–147419. https://doi.org/10.1109/ACCESS.2019.2946515
21. Bruynseels, K., Santoni de Sio, F. and van den Hoven, J. Digital Twins in health care: ethical implications of an emerging engineering paradigm. *Front. Genet.*, 2018, **9**, 31. https://doi.org/10.3389/fgene.2018.00031
22. Terkaj, W., Tolio, T. and Urgo, M. A virtual factory approach for in situ simulation to support production and maintenance planning. *CIRP Ann.*, 2015, **64**(1), 451−454. https://doi.org/10.1016/j.cirp.2015.04.121
23. Tšukrejev, P., Kruuser, K. and Karjust, K. Production monitoring system development for manufacturing processes of photovoltaic modules. *Proc. Est. Acad. Sci.*, 2019, **68**(4), 401−406. https://doi.org/10.3176/proc.2019.4.09
24. Graessler, I. and Poehler, A. Integration of a digital twin as human representation in a scheduling procedure of a cyber-physical production system. In *Proceedings of the 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Singapore, December 10−13, 2017*. IEEE, 2018, 289−293. https://doi.org/10.1109/IEEM.2017.8289898
25. Wang, Q., Jiao, W., Wang P. and Zhang, Y. Digital Twin for human-robot interactive welding and welder behavior analysis. *IEEE/CAA J. Autom. Sin.*, 2021, **8**(2), 334−343. https://doi.org/10.1109/JAS.2020.1003518
26. Buldakova, T. I. and Suyatinov, S. I. Hierarchy of human operator models for Digital Twin. In *Proceedings of the International Russian Automation Conference (RusAutoCon), Sochi, Russia, September 8−14, 2019*. IEEE, 1−5. https://doi.org/ 10.1109/RUSAUTOCON.2019.8867602
27. Milgram, P., Drascic, D., Grodski, J. and Restogi, A. Merging real and virtual worlds. *Proc. IMAGINA'95*, 1995.
28. Kuts, V., Sarkans, M., Otto, T., Tähemaa, T. and Bondarenko, Y. Digital Twin: concept of hybrid programming for industrial robots − use case. In *Proceedings of the ASME 2019 International Mechanical Engineering Congress and Exposition-Advanced Manufacturing*. 2020, **2B**. https://doi.org/10.1115/IMECE2019-10583
29. Horváth, G. and Erdős, G. Gesture control of cyber physical systems. *Procedia CIRP*, 2017, **63**, 184−188. https://doi.org/10.1016/j.procir.2017.03.312
30. Pérez, L., Rodríguez-Jiménez, S., Rodríguez, N., Usamentiaga, R. and García, D. F. Digital Twin and virtual reality based methodology for multi-robot manufacturing cell commissioning. *Appl. Sci.*, 2020, **10**(10), 3633. https://doi.org/10.3390/app10103633
31. Vosniakos, G. C., Ouillon, L. and Matsas, E. Exploration of two safety strategies in human-robot collaborative manufacturing using Virtual Reality. *Procedia Manuf.*, 2019, **38**, 524−531. https://doi.org/10.1016/j.promfg.2020.01.066
32. Peruzzini, M., Grandi, F. and Pellicciari, M. Exploring the potential of Operator 4.0 interface and monitoring. *Comput. Ind. Eng.*, 2020, **139**, 105600. https://doi.org/10.1016/j.cie.2018.12.047
33. Concannon, D., Flynn, R. and Murray, N. A quality of experience evaluation system and research challenges for networked virtual reality-based teleoperation applications. In *Proceedings of the 11th ACM Workshop on Immersive Mixed and Virtual Environment Systems − MMVE*. 2019, 10−12. https://doi.org/10.1145/3304113.3326119
34. Whitney, D., Rosen, E., Phillips, E., Konidaris, G. and Tellex, S. Comparing robot grasping teleoperation across desktop and virtual reality with ROS reality. In *Robotics Research. Springer Proceedings in Advanced Robotics*. Springer, Cham, 2020, **10**, 335−350. https://doi.org/10.1007/978-3-030-28619-4_28
35. ISO 9241-210:2019. Ergonomics of Human-System Interaction − Human-Centred Design for Interactive Systems. International Organization for Standardization, Geneva, 2010. https://www.iso.org/standard/77520.html

36. Sell, R., Leier, M., Rassõlkin, A. and Ernits, J.-P. Self-driving car ISEAUTO for research and education. In *Proceedings of the 2018 19th International Conference on Research and Education in Mechatronics, Delft, Netherlands, June 7–8, 2018*. IEEE, 111–116. https://doi.org/10.1109/REM.2018.8421793
37. Kuts, V., Otto, T., Tähemaa, T. and Bondarenko, Y. Digital twin based synchronised control and simulation of the industrial robotic cell using virtual reality. *J. Mach. Eng.*, 2019, **19**(1), 128–144. https://doi.org/10.5604/01.3001.0013.0464
38. Rebelo, F., Noriega, P., Duarte, E. and Soares, M. Using virtual reality to assess user experience. *Hum. Factors*, 2012, **54**(6), 964–982. https://doi.org/10.1177/0018720812465006

Kasutajakeskne disain tööstuslikes koostööautomaatsüsteemides

Simone Luca Pizzagalli, Vladimir Kuts ja Tauno Otto

Autonoomsed süsteemid ja koostöörobootika on Tööstus 4.0 paradigma ühed olulisemad tugitehnoloogiad. Siia hulka kuuluvad täiustatud simulatsioonid, digitaalsed kaksikud ja uudsed inim-masinliidesed. Nende tehnoloogiate areng koos kohandatud tootmisprotsessidele esitatavate kõrgemate nõuetega eeldab tihedamat koostööd operaatorite ja automatiseeritud süsteemide vahel. Seetõttu tuleb ümber defineerida kogu senine töökorraldus: kuidas inimoperaatorid masinaid haldavad ja nendega suhtlevad ning kuidas neid selleks kohandatavad liidesed, simulatsioonid ja reaalajas andmete kogumine ning analüüs toetavad. Uued inimese ja roboti koostöö paradigmad on ülimalt olulised stsenaariumis, kus piirid inimeste ja masinate sooritatavate ülesannete vahel on paindlikud ja hägusad. Standardite, disainimeetodite, programmeerimisliideste ja hindamismeetodite uuesti määratlemine on nende tehnoloogiliste ja tootmismuudatuste hõlbustamisel kesksel kohal. Inimvõimekuse suurendamine töökohal tingib vajaduse määratleda nõuete raamistik, mis integreeriks inim-, organisatsiooni- ja tootmisvajadused samasse stsenaariumi ja töövoogu. Antud uurimus pakub lahendusena välja kasutajakeskse disaini eelistamise, mis on inimese ja roboti koostöösüsteemidele avatud väljakutsete lahendamisel ülioluline. Meie metoodika peab digitaalseid kaksikuid ning liit- ja virtuaalreaalsuse tehnoloogiaid peamisteks tööriistadeks nüüdisaegsete koostööstsenaariumide kujundamisel, juhtimisel ja hindamisel.

II publication

II

S.L. Pizzagalli, V. Kuts, Y. Bondarenko, and T. Otto. Evaluation of virtual reality interface interaction methods for digital twin industrial robot programming and control, a pilot study. In *ASME International Mechanical Engineering Congress and Exposition*, volume 85567, page V02BT02A005. American Society of Mechanical Engineers, 2021

Proceedings of the ASME 2021
International Mechanical Engineering Congress and Exposition
IMECE2021
November 1-4, 2021, Virtual, Online

IMECE2021-69408

EVALUATION OF VIRTUAL REALITY INTERFACE INTERACTION METHODS FOR DIGITAL TWIN INDUSTRIAL ROBOT PROGRAMMING AND CONTROL, A PILOT STUDY

Simone Pizzagalli[1] **Vladimir Kuts[1]** **Yevhen Bondarenko[1]** **Tauno Otto[1]**
[1]Tallinn University of Technology, Tallinn, Estonia
Department of mechanical and industrial engineering

ABSTRACT

Extended Reality (XR) interfaces for Human-Robot Interaction (HRI) allow for safe and flexible control and programming of advanced industrial systems and integrate the operator as part of the Digital Twins (DT) interfaces of real collaborative systems. These technologies already show their potential in real-life scenarios by providing valuable training, quality control, maintenance, remote control and assistance tools. XR could also favor the development of new metrics and methods in human-robot collaboration (HRC) by providing safe and reproducible testbeds for the standardization of HRI in collaborative scenarios. This can be achieved by attaining a reliability level that would avoid the degradation of user experience in HRI. XR DT systems and interfaces need to be assessed on their impact on the operator, but also in terms of interaction efficacy and efficiency within the virtual scenario. This evaluation should be performed on both virtual reality (VR) visualization and interaction hardware and virtual user interfaces (VUI) in immersive scenarios. There is a lack of evaluation of performance and usability of VUI input methods. Our work aims at assessing three types of input paradigms with a VUI controlling a DT robot arm in a simple pick and place task and which can be based on different VR controller devices. Objective metrics are collected during the task while NASA TLX, SUS and UEQ are used to assess user workload, usability of the system and user experience for each interaction method.

Keywords: human-robot interaction, virtual interfaces, interaction methods, digital twins

1. INTRODUCTION

Adaptable, intuitive, and efficient user interfaces (UI) are crucial elements for industrial process digitalization within the Industry 4.0 framework. UIs are essential components in industrial robot programming and control, especially in modern collaborative scenarios requiring fast robot reprogramming, production changeover, and reliable safety methods [1]. Recent works demonstrate the strength of digital twin (DT) extended reality (XR) interfaces in providing flexible methods for safety evaluation [2], workspace design and assessment, programming [3], training, online collaboration [4], and teleoperation [5]. By including the operator in the DT of the physical industrial system, these interfaces provide a powerful means of interaction and constitute a workbench for system efficiency, human-robot collaboration (HRC) assessment, user evaluation and monitoring towards the operator 4.0 paradigm [6]. Nonetheless, XR systems are likely to increase user stress levels, eye fatigue, and sickness [7]. Applications in this field tend to focalize on the overall DT systems performance and evaluation, considering virtual user interface (VUI) interaction paradigms barely as a means to an end. A physical interface, usually a tech pendant, mediates traditional robotic arm control and programming. Interactions within the DT Virtual Reality (VR) scenario are generally mediated by the visualization and control hardware and devices. The interaction with the VUI can be delivered in different ways, depending on both device capabilities, VR software implementation and adopted interaction metaphors. These can highly influence user task performance, experience, and workload while interacting with the DT or VUI. Despite being designed with different physical and morphological characteristics, depending on the producer, VR controllers have similar button mapping which can be easily interchanged. In our previous work, based on the industrial robot DT VUI presented in [8], we investigated the efficiency of DT XR interfaces in robot control by comparing user task performance with a physical teach pendant. Based on previous research results, this work aims at the comparison and evaluation of VUI interaction methods based on different input system design approaches. This study proposes a virtual interface interaction task implemented in Unity3D software. Interaction with the VUI components for robot control and manipulation are provided by three distinct input methods, which present a scaled level of abstraction from hardware-based to hand avatar interaction. In particular, the user interacts with the VUI by using VR device controller button mapping as direct input, traditional VR pointer UI selection and interaction, and physical-based manipulation via user hand avatar. The study aims at evaluating the three different methods while being indifferent to the specific VR controller devices. The goal is to assess the most efficient and usable input method for VR-based UIs and provide a starting point for further design, development and, evaluation of XR DT interfaces for HRC, programming, and control.

1.1 Related works

The implementation of XR technologies in the industrial field and the possibility of using immersive DT environments as interfaces for advanced HRC scenarios pushed recent studies towards the efficiency assessment of VUI in robot control and interaction methods with the DTs [9]. Several publications approach robot operation through natural user interfaces (NUI), such as gesture control [10], or tangible user interfaces (TUI). The study presented in [11] compares TUI robot operation with a traditional keypad interface. The study in [12] evaluates the efficiency of robot grasping by comparing keyboard and monitor with hand tracking manipulation both in a screen-based or immersive scenario. The work in [13] presents an experimental comparison of two types of VR controller-based interaction methods aimed at robot operation but without employing any VUI. Several other works present experimental assessment of VR interaction modalities, input paradigms and devices in different research fields. An example of comparative performance assessment of different interaction methods with virtual graphs and VUI can be found in [14]. Results show that traditional keyboard and mouse configuration and gamepad configuration, also in combination with VR visualization, give better results than hand tracking and NUI in immersive test settings. Similar results are confirmed by the study presented in [15]. A common remark about direct VUI input using hand tracking is the reliability of the hardware tracking system and the familiarity of the user with this technology. The results of the experimental study presented in [16] show that gesture-based interactions for VUI are less efficient than point and click and controller-based interactions. The comparison of traditional gamepad devices and virtual reality controllers does not offer clear positive performance results in favor of the latter [17]. The study in [18] presents a comparative evaluation of different user interface devices for a VR mining application. A comparison between VUI interaction methods based on trackpad and raycast pointer input is presented in [19]. While comparing input methods based on VR controller hardware without introducing further interaction modalities the study implements two different interface designs and positions in the virtual environment (VE). The experiment presented in [20] proposes a comparison between different input methods for text entry in virtual reality. The study presents several input paradigms such as controller pointing, head pointing, controller tapping, freehand and discrete or continuous cursor focalizing on the keyboard interface as a standard method for text entry in computer systems and VR environments. While discrete and continuous cursor input performed better in terms of word per minute the free hand and controller pointing showed better results in terms of user experience. Our work aims at the assessment of VUI controller-based input methods for industrial robot control and manipulation without introducing different interaction modalities and being indifferent to specific VR controller hardware.

2. MATERIALS AND METHODS

This study is based on the DT VR control interface presented in [8]. The system will be extensively described in the following paragraphs. In our previous work, we tested the VUI against traditional programming teach pendant assessing user experience (UX) degradation in XR by evaluating task and system-based objective metrics and measuring the overall impact on the operator by means of quantitative and qualitative data analysis. This work aims at introducing a further and more specific evaluation by detecting metrics related to the user experience of VUI input and navigation paradigms in a simple robot pick and place task.

2.1 Digital Twin Environment and User Interface

The DT system used in this experimental work is developed in the TalTech Industrial Virtual and Augmented Reality laboratory (IVAR). The VE includes several digital couples of real laboratory industrial systems and is shown in Figure 1.

FIGURE 1: THE IVAR LAB DIGITAL TWIN VIRTUAL ENVIRONMENT

The digital replica is developed by using Unity game engine and additional scripting logic developed in C# programming language for control and real-world synchronization. This study focalizes on the Yaskawa Motoman GP8 robotic arm shown in Figure. 2.

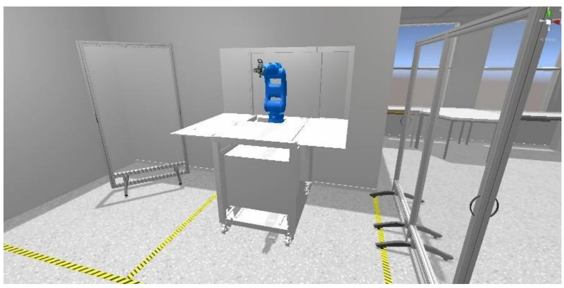

FIGURE 2: THE MOTOMAN GP8 DIGITAL TWIN

The virtual robot replicates the kinematics, joint controls and limits of its real counterpart. Additionally, it can be manipulated in two modes being either coupled with the real machine or only manipulated in the VE. The system allows for directly teleoperating the real or virtual robot or creating a task program stored for later use. As we do not focalize on twinning and communication performance, in this study and our experiment we employ the virtual mode with direct robot manipulation.

The world space UI presented in Figure 3. is dedicated to the control of the robot twin and positioned in front of it in the VE.

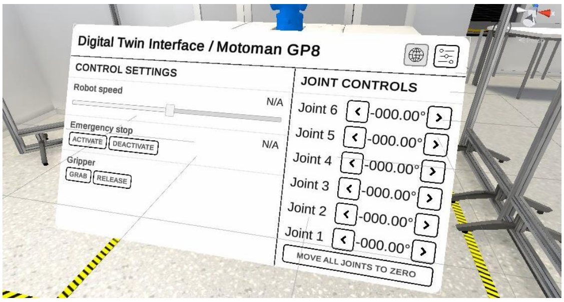

FIGURE 3: THE MOTOMAN GP8 VIRTUAL USER INTERFACE

The interface comprises several buttons, toggles and controls grouped by functions. A header allows access to a dedicated programming and coupled-mode connection interface. The robot control interface (RCUI) is divided in control settings including emergency stop activation and deactivation, robot speed slider, and the robot gripper control buttons. On the right side of the RCUI each robot joint can be controlled by several dedicated buttons. The rotation angle for each joint is visualized in the interfaces while visual textual feedbacks are given for the related speed, emergency stop, and gripper control functions.

2.2 Equipment

The VR application runs on a laptop computer mounting a Intel core I7-6700HQ, Nvidia GEFORCE GTX 970m graphic processing unit and 16 GB of ram. The user involved in the experiments navigates and interacts with the VE through Oculus Rift S head-mounted display (HMD) and dedicated controllers. The headset and controllers employed during the experiment are shown in Figure 4.

FIGURE 4: HEAD MOUNTED DISPLAY AND CONTROLLERS EMPLOYED DURING THE EXPERIMENTS

Before the experimental sessions, the system is calibrated once by defining the user working area and floor height in the VE. Furthermore, each user is provided with a disposable HMD display paper cover to avoid direct contact with the face. The experiments are performed in the IVAR lab at Tallinn University of Technology.

2.3 Proposed Interaction methods

This study aims at determining the user experience and VRUI interaction performance during a DT robotic arm pick and place task. We selected and implemented three popular controller-based UI interaction paradigms aiming at refining our previous experimental work results and moving forward in a step-by-step assessment of VUI interaction methods and user experience.

The first method adopts a controller-based UI navigation and selection model (CBI) and is implemented by adopting the Unity Input System v1.0.2. This software package allows for the creation of a set of action maps associated to specific VR scene UI or Player controls. Each action can be bound to one or more specific input values regardless to the input hardware. This allows for a fast and efficient input binding and the creation of flexible plug-and-play inputs based on different hardware systems (keyboard, XR controllers, game controllers etc.). By adopting this specific method, the software is able to map each VUI action (e.g., pointer click or scroll) to any of the most popular XR controllers which, in any case, are offering similar design, button mappings and positions. This specific VR interaction method works similarly to the old-fashioned monitor desktop-based video game UI navigation and selection. The VUI elements are highlighted upon selection. The user can select different UI interactable elements and browse through them by means of the controller joystick. By pressing the controller trigger the UI elements are clicked or interacted with. Oculus Rift S controller trigger and joystick are shown in Figure 5.

FIGURE 5: OCULUS CONTROLLER TRIGGER AND JOYSTICK

The physical interactions with the controller buttons are kept to a minimum to avoid confusion and allow for an easier VUI selection and control process. The specific Input System binding structure adopted for this experiment is shown in Table 1.

TABLE 1: INPUT MAPPING FOR XR CONTROLLERS

Action	Controller binding
UI navigation	Right or Left XR controller Joystick
Click	Right or Left trigger button pressed
Slider interaction	Right or Left XR controller Joystick (left and right only)

The second method is based on ray cast pointer selection (RCPI). In this scenario, the pointer ray cast is automatically triggered and visible when the user directs the controller at any UI element in the VE. Once the pointer hits any interactable element in the UI (e.g., buttons, sliders, or toggles) this is highlighted, and it is possible to interact with it by pressing the controller trigger button. The source of the ray casted pointer is positioned on top of the virtual avatar controller. Each controller, right and left hand can be used interchangeably. The color of the ray will change to blue when the controller input becomes dominant. Figure 6 shows an example of ray cast pointer button selection.

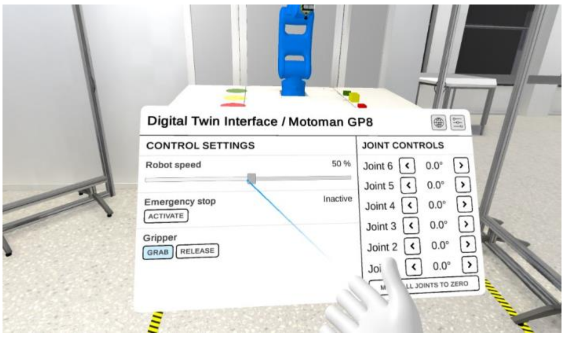

FIGURE 6: RAYCAST BASED VRUI BUTTON SELECTION AND INTERACTION (RCPI)

The third interaction method allows the user to manipulate the UI elements directly with his avatar hands (DI). The user is still holding the XR controller which enables hand position and rotation tracking in the VE. The interaction with the VUI grounds on a modified version of the ray cast pointer selection method presented before. In this case the ray cast pointer is not visible and attached to the hand avatar index fingertip. A specific script logic allows the pointer to be triggered when the finger is directed towards a UI in the VE. Furthermore, the pointer is always directed perpendicularly to the interface elements. The interaction logic allows the system to detect the distance between the fingertip and UI interactable elements. Once the distance is less than five millimeters the UI element is selected and changes color. Pressing a button or moving a scrollbar handle happens when the hand avatar fingertip distance to the UI elements is less than a predefined threshold. This method attempts in replicating in VR a real-world interaction with a touch screen interface. Figure 7 shows an example of the direct VUI interaction method.

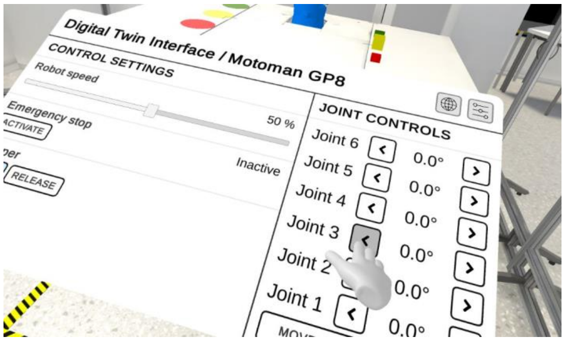

FIGURE 7: DIRECT VR UI INTERACTION METHOD (DI)

2.4 Tasks and study protocol

The task proposed in this study is a simple pick and place assignment similar to the one described in our previous experiments and making use of the VR DT system at TalTech IVAR lab. The user should relocate three cubes of different sizes and colors to a specific target position on top of the robot's workbench. The target positions are highlighted with colors similar to the cubes namely red, green and yellow. The VUI navigation and interaction tasks that each subject should perform during the experiment are presented in Table 2.

TABLE 2: VRUI NAVIGATION AND INTERACTION TASKS

UI interaction protocol
Access the robot interface
Deactivate the emergency stop option
Move all joints to zero
Increase the robot speed to 60%
Interact with the joints controls to grab the first object
Grab the first object
Interact with the joints controls to move the first object
Release the first object in place
Interact with the joints controls to grab the second object
Grab the second object
Interact with the joints controls to move the second object
Interact with the joints controls to grab the third object
Grab the third object
Interact with the joints controls to move the third object
Release the third object in place
Move all joints to zero

Each subject is required to perform the task three times by navigating and interacting with the VUI by means of the previously presented CBI, RCPI, and DI methods. To be able to counterbalance the results the subjects taking part in the experiment are divided into 3 subgroups. Each group performs the required task with a different sequence of interaction methods. Figure 8 shows the interaction method protocol for each subgroup in each session.

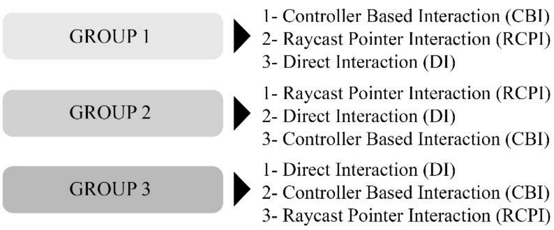

FIGURE 8: INTERACTION METHOD SEQUENCING FOR EACH SUBJECT SUBGROUP

Before each session, the users are introduced to both the main experiment objectives and task (object pick and place) and the VRUI navigation protocol. The type of VUI interactions and the specific controls required in each session are also fully introduced before the test. Regardless of the familiarity with VR navigation and equipment each user is given 2 minutes to familiarize with the HMD and controllers and explore the Oculus VR default setup environment. After each session, the user is presented with three questionnaires evaluating workload, usability of the system, and user experience for each interaction method. Figure 9 shows the experimental protocol for each user.

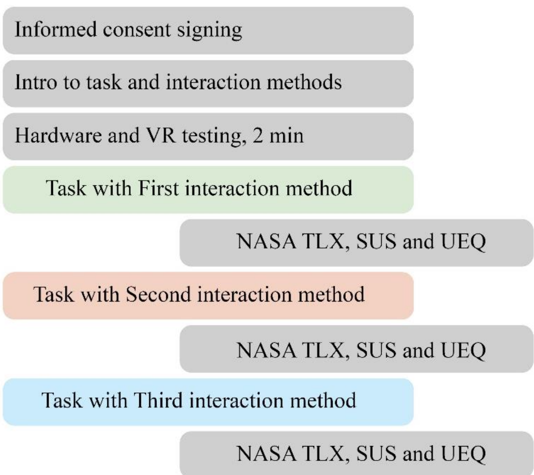

FIGURE 9: EXPERIMENTAL PROTOCOL FOR EACH USER

At the beginning of the session, the user is located in front of the VUI without the possibility of teleporting or moving in the environment. This is solution is adopted to avoid teleporting away from the interface especially while performing tests with the CBI interaction method.

2.5 Participants

15 subjects of different age and gender are recruited among students and employees of the faculty of engineering at Tallinn University of Technology. The group is divided into 3 subgroups of 5 participants each. Each subgroup undergoes the test starting from a different interaction method as explained in the previous paragraph. Due to the current pandemic and ongoing restrictions in Estonia forcing full online work and studies starting from the beginning of March till the end of the academic semester the experiment is performed by 6 subjects only. The same protocol, tasks, and interaction methods are maintained but only a basic statistical analysis is performed due to the limited collected data. The subjects grouping and counterbalancing is maintained. Each participant signed an informed consent on data usage and handling and filled an introductory questionnaire on demographics including questions on their scholastic level, gender, age, familiarity with VR applications, familiarity with immersive VR equipment and controllers, familiarity with robot control and programming, use of gaming consoles and videogames in general. Familiarity with VR applications, equipment, and robot control or programming was assessed by using a 5-point Likert scale.

2.6 Metrics

Qualitative and quantitative data is collected during and after each session. Quantitative data focus on task completion performance, therefore time spent in completing the pick and place assignment for all the cubes is saved for each type of interaction method. Qualitative data is mainly focusing on task workload, usability, and user experience during the robot manipulation in VR. For this purpose, we employ three different validated questionnaires. NASA Task Load Index (TLX) [21] is used to assess workload during the task performance for each user with each different interaction method. The System Usability Scale (SUS) [22] is adopted to detect the general usability of the system while the User Experience Questionnaire (UEQ) [23] is assessing the user experience while interacting with the DT VUI. A think aloud protocol is employed to record users' comments and remarks while performing the tasks. The comments are annotated during each experimental session by one of the researchers involved in the study.

3. RESULTS AND DISCUSSION

3.1 Participants

A total of 4 male and 2 female participants with average age of 32 years old (SD ± 9) took part in the study. Three subjects hold a title of Doctor of Philosophy, one has a Bachelor Diploma and two a Master of Science diploma. Most of the subjects declared of having previous experiences with both AR/VR hardware and software and none of them was completely unaware of game consoles and controller's usage. Two subjects declared not being familiar with HMD and controllers and only one disagreed in having experience with AR/VR applications. Experience and familiarity with robot control and programming gave more diversified results with samples being almost evenly distributed along the measuring scale. Two subjects have clearly extensive experience with industrial robotic systems. A summary of the experimental group characteristics and demographic data are presented in Table 3.

TABLE 3: PARTICIPANTS DEMOGRAPHICS AND SKILLS

Age	32 ± 9
Male	4
Female	2
Education level	
• High School	0
• Bachelor Diploma	1
• Master Diploma	2
• PhD	3
Users have previous experiences with HMD and controllers	
• Strongly agree	0
• Agree	4
• Neutral	0
• Disagree	1
• Strongly disagree	1
Users have previous experiences with AR VR applications	
• Strongly agree	0
• Agree	4
• Neutral	1
• Disagree	0
• Strongly disagree	1
Users have previous experiences with gaming consoles	
• Strongly agree	2
• Agree	3
• Neutral	1
• Disagree	0
• Strongly disagree	0
Users have previous experiences in robot control and programming	
• Strongly agree	2
• Agree	1
• Neutral	1
• Disagree	1
• Strongly disagree	1

3.2 Task performance

Results in task execution timing show that the most efficient interaction method in terms of average total time in relocating the three cubes is RCPI with a value of 195,50 seconds (SD 127,19 s). This method is followed by CBI with an average time value of 312,33 s (SD 115,5 s) and DI with 351,17 seconds (SD 138,16). A summary of the average timing in completing the task per interaction method is given in Table 4. Times are reported in seconds.

TABLE 4: AVARAGE TIME TO COMPLETE THE TASK FOR EACH INTERACTION METHOD

Method	Mean in seconds	SD in seconds
RCPI	195.50	± 127.19
CBI	312.33	± 115.50
DI	351.17	± 138.16

3.3 Workload, system usability and user experience

Table 5 reports the raw unweighted data results from the Nasa TLX questionnaire.

TABLE 5: NASA TLX QUESTIONNAIRE RAW SCORES COMPARISON FOR EACH METHOD

	RCPI	CBI	DI
Mental	25.00	42.50	42.50
Physical	12.50	17.50	33.33
Temporal	18.33	27.50	44.17
Performance	24.17	30.83	33.33
Effort	20.00	18.33	34.17
Frustration	11.67	20.83	33.33
Overall	18.61	26.25	36.81

The table rows report the values for the questionnaire metrics regarding mental, physical, and temporal demand, performance, effort and frustration for each interaction method. The last raw reports the overall mean value score for each specific method. Results show generally low scores for each type of interaction. The best overall score is associated with the RCPI method with an overall value of 18.61. This result is followed by CBI, with a score of 26.25, and DI, 36.81. CBI method shows higher values for mental demand, 42.50, temporal demand, 27.50 and performance, 30.83. The direct interaction UI method has generally higher values for each metric with two maximum values for mental demand, 42.50, and temporal demand, 44.17.

SUS questionnaire results show that the RCPI has optimal usability with an average value of 90.83 (SD ±7.31). Usability scores for the CBI and DI methods are similar with average values of 72.91 (SD ±15.97) the former and 70.83 (SD ±12.72) the latter which, in any case, indicate good usability of the system. Table 6 shows the results for the SUS questionnaire for each interaction method.

TABLE 6: SUS QUESTIONNAIRE RESULTS

	Mean	SD
RCPI	90.83	±7.31
CBI	72.91	±15.96
DI	70.83	±12.72

The user experience questionnaire shows good results both for the RCPI and DI interaction methods with values above average (>0.8) for both pragmatic and hedonic quality of experience. RCPI and DI methods have overall mean score values of 2.06 and 1.25 respectively. The CBI gave neutral results for the hedonic quality and an overall score of 1.15. UEQ results are reported in table 7.

TABLE 7: UEQ RESULTS

	Pragmatic	Hedonic	Overall
RCPI	2.46	1.66	2.06
CBI	1.67	0.65	1.15
DI	1.25	1.25	1.25

3.5 User comments and observations

The speak aloud protocol allows to detect direct impressions from the user while performing the task with each interaction methods. Regarding the CBI, users commented that the method seems a bit outdated but somehow easier to handle. One user commented that this specific method showed a higher responsiveness and speed of the robot joint movements. Another subject stated that the RCPI was the most comfortable interaction method while two different users noticed that DI was too sensitive and physically demanding because the user have to physically move close to the interface and change focus of attention from the button to the robots for each interaction. This is also noticed by the investigators during task performance observations. General comments on the VUI highlighted the necessity for clearer feedbacks of the system both regarding the grabbing and release function and about the position of the robot joints. A user commented that an explicit reference system would allow to understand which joint is moving while pressing a specific button and in which direction. One comment addressed the joint control reset button being too close to the joint control buttons. Another user had problems with Oculus lenses getting 'foggy' and occluding the view of the VE.

3.6 Discussion

The experiment shows overall good results in terms of usability and user experience for each presented control methods. The best interaction method overall is the RCPI both in terms of timing in completing the task, low user workload, usability, and user experience. Nevertheless, the RCPI method standard deviation value for task completion performance is quite high. CBI gave good results in terms of performance. There is no relevant difference in time results for this method and the DI although, in the case of CBI, standard deviation presents a smaller value. System usability for both methods is good and gave similar score results. DI method performed better in terms of user experience with CBI having low scores in hedonic quality of the system. The interface in general needs more feedbacks and should be more clearly showing which joint each button is controlling and in which direction. The DI method seems to offer a natural way of interacting with VUI as in real world mobile based touch interfaces. Nasa TLX shows that this method is relevantly more demanding almost in all metrics detected by the questionnaire. Mental and temporal workload values are especially high probably because the interaction is more dependable on the VUI position and size forcing the user to move a lot and change the focus of attention from the buttons to the robot at each interaction. The CBI trigger button sensitivity should be also addressed as it resulted in the perception of the robot moving faster than expected. A feedback on the robot gripper activation area and grab and release button actions should also be implemented. Observations during task performance showed that some of the users focused on precisely matching gripper orientation with the cubes while this was unnecessary to achieve a successful grab action. This, in some case, might have influenced time performance in completing the task.

The presented study shows many limitations starting from the number of subjects involved in the experiments. Due to the COVID19 restrictions at Tallinn University of Technology it was difficult to recruit more users. The lack of data results in poor statistical analysis and unreliable system usability and user experience questionnaires results. Also, some improvements on the system and interface are necessary to achieve better experiment results and a more efficient robot control.

Future works will consider testing the system on a larger experimental group, fixing some interaction and system flaws and improve interface accessibility especially in DI method. Precision of the trigger button click should be improved in the CBI method. Clearer feedbacks on robot position and joint control movement in relation with the VUI buttons should be provided to avoid unnecessary workload on the user in any of the interaction methods. The VUI can be partially redesigned and resized to allow a better accessibility for the DI method. Avatar hands might be as well improved. Further developments and user studies might include comparison of different interface layout and implementation of different interaction methods.

4. CONCLUSION

Despite the limitations, the experiments presented in this study give positive results in terms of VUI usability and user experience in a DT VR human robot interaction task. Further user studies should focalize in comparing interaction methods with VUI for robot control, programming and HRC to achieve usable, efficient DT VR interfaces and assessing novel metrics and standards for HRI.

ACKNOWLEDGEMENTS

Our acknowledgements go to the Estonian Research Council project TT2 "Smart Industry Centre", and everyone who participated in the study.

REFERENCES

[1] V. Villani, F. Pini, F. Leali, and C. Secchi, "Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications," *Mechatronics*, vol. 55, no. February, pp. 248–266, 2018, doi: 10.1016/j.mechatronics.2018.02.009.

[2] G.-C. Vosniakos, L. Ouillon, and E. Matsas, "Exploration of two safety strategies in human-robot collaborative manufacturing using Virtual Reality," Procedia Manuf., vol. 38, no. 2019, pp. 524–531, 2019, doi: 10.1016/j.promfg.2020.01.066.

[3] S. K. Ong, A. W. W. Yew, N. K. Thanigaivel, and A. Y. C. Nee, "Augmented reality-assisted robot programming system for industrial applications," *Robot. Comput. Integr. Manuf.*, vol. 61, no. May 2019, p. 101820, Feb. 2020, doi: 10.1016/j.rcim.2019.101820.

[4] V. Kuts, T. Otto, Y. Bondarenko, and F. Yu, “Digital twin: Collaborative virtual reality environment for multi-purpose industrial applications,” in *ASME International Mechanical Engineering Congress and Exposition, Proceedings (IMECE)*, 2020, vol. 2B-2020, doi: 10.1115/IMECE2020-23390.

[5] G. A. Yashin, D. Trinitatova, R. T. Agishev, R. Ibrahimov, and D. Tsetserukou, “AeroVr: Virtual Reality-based Teleoperation with Tactile Feedback for Aerial Manipulation,” in *2019 19th International Conference on Advanced Robotics (ICAR)*, Dec. 2019, pp. 767–772, doi: 10.1109/ICAR46387.2019.8981574.

[6] D. Romero, P. Bernus, O. Noran, J. Stahre, and Å. Fast-Berglund, “The Operator 4.0: Human Cyber-Physical Systems & Adaptive Automation Towards Human-Automation Symbiosis Work Systems,” in *IFIP Advances in Information and Communication Technology*, no. September, 2016, pp. 677–686.

[7] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson, “Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems,” *Displays*, vol. 29, no. 2, pp. 58–69, 2008, doi: 10.1016/j.displa.2007.09.005.

[8] V. Kuts, M. Sarkans, T. Otto, T. Tähemaa, and Y. Bondarenko, “Digital Twin: Concept of Hybrid Programming for Industrial Robots — Use Case,” in *Volume 2B: Advanced Manufacturing*, Nov. 2019, vol. 2B-2019, doi: 10.1115/IMECE2019-10583.

[9] J. O. Oyekan *et al.*, “The effectiveness of virtual environments in developing collaborative strategies between industrial robots and humans,” *Robot. Comput. Integr. Manuf.*, vol. 55, no. July 2018, pp. 41–54, 2019, doi: 10.1016/j.rcim.2018.07.006.

[10] G. Horváth and G. Erdős, “Gesture Control of Cyber Physical Systems,” *Procedia CIRP*, vol. 63, pp. 184–188, 2017, doi: 10.1016/j.procir.2017.03.312.

[11] C. Guo and E. Sharlin, “Utilizing physical objects and metaphors for human robot interaction,” 2008.

[12] D. Whitney, E. Rosen, E. Phillips, G. Konidaris, and S. Tellex, “Comparing Robot Grasping Teleoperation Across Desktop and Virtual Reality with ROS Reality,” no. February, 2020, pp. 335–350.

[13] R. Hetrick, N. Amerson, B. Kim, E. Rosen, E. J. D. Visser, and E. Phillips, “Comparing Virtual Reality Interfaces for the Teleoperation of Robots,” *2020 Syst. Inf. Eng. Des. Symp. SIEDS 2020*, 2020, doi: 10.1109/SIEDS49339.2020.9106630.

[14] U. Erra, D. Malandrino, and L. Pepe, “Virtual Reality Interfaces for Interacting with Three-Dimensional Graphs,” *Int. J. Hum. Comput. Interact.*, vol. 35, no. 1, pp. 75–88, 2019, doi: 10.1080/10447318.2018.1429061.

[15] A. Santos-Torres, T. Zarraonandia, P. Díaz, T. Onorati, and I. Aedo, “An empirical comparison of interaction styles for map interfaces in immersive virtual environments,” *Multimed. Tools Appl.*, vol. 79, no. 47–48, pp. 35717–35738, 2020, doi: 10.1007/s11042-020-08709-9.

[16] H. Kharoub, M. Lataifeh, and N. Ahmed, “3D user interface design and usability for immersive VR,” *Appl. Sci.*, vol. 9, no. 22, 2019, doi: 10.3390/app9224861.

[17] M. Ali and R. E. Cardona-Rivera, “Comparing Gamepad and Naturally-mapped Controller Effects on Perceived Virtual Reality Experiences,” *Proc. - SAP 2020 ACM Symp. Appl. Percept.*, 2020, doi: 10.1145/3385955.3407923.

[18] H. Kim and Y. Choi, “Performance comparison of user interface devices for controlling mining software in virtual reality environments,” *Appl. Sci.*, vol. 9, no. 13, 2019, doi: 10.3390/app9132584.

[19] P. Monteiro, H. Coelho, G. Gonçalves, M. Melo, and M. Bessa, “Comparison of radial and panel menus in virtual reality,” *IEEE Access*, vol. 7, pp. 116370–116379, 2019, doi: 10.1109/ACCESS.2019.2933055.

[20] M. Speicher, A. M. Feit, P. Ziegler, and A. Krüger, “Selection-based text entry In Virtual Reality,” *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2018-April, no. May, 2018, doi: 10.1145/3173574.3174221.

[21] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research,” *Adv. Psychol.*, vol. 52, no. C, pp. 139–183, 1988, doi: 10.1016/S0166-4115(08)62386-9.

[22] J.Brooke, “SUS: A ‘Quick and Dirty’ Usability Scale,” in *Usability Evaluation In Industry*, vol. 15, no. 8, CRC Press, 1996, pp. 207–212.

[23] B. Laugwitz, T. Held, and M. Schrepp, “Construction and evaluation of a user experience questionnaire,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2008, vol. 5298 LNCS, pp. 63–76, doi: 10.1007/978-3-540-89350-9_6.

III publication

III

V. Kuts, J. A. Marvel, M. Aksu, **S.L. Pizzagalli**, M. Sarkans, Y. Bondarenko, and T. Otto. Digital twin as industrial robots manipulation validation tool. *Robotics*, 11(5):113, 2022

robotics

MDPI

Article

Digital Twin as Industrial Robots Manipulation Validation Tool

Vladimir Kuts [1,*], Jeremy A. Marvel [2], Murat Aksu [2], Simone L. Pizzagalli [1], Martinš Sarkans [1], Yevhen Bondarenko [1] and Tauno Otto [1]

1 Department of Mechanical and Industrial Engineering, Tallinn University of Technology, 19086 Tallinn, Estonia
2 National Institute of Standards and Technology, Gaithersburg, MD 20899, USA
* Correspondence: vladimir.kuts@taltech.ee

Abstract: The adoption of Digital Twin (DT) solutions for industrial purposes is increasing among small- and medium-sized enterprises and is already being integrated into many large-scale companies. As there is an increasing need for faster production and shortening of the learning curve for new emerging technologies, Virtual Reality (VR) interfaces for enterprise manufacturing DTs seem to be a good solution. Furthermore, with the emergence of Industry 5.0 (I5.0) paradigm, human operators will be increasingly integrated in the systems interfaces though advanced interactions, pervasive sensors, real time tracking and data acquisition. This scenario is especially relevant in collaborative automated systems where the introduction of immersive VR interfaces based on production cell DTs might provide a solution for the integration of the human factors in the modern industrial scenarios. This study presents experimental results of the comparison between users controlling a physical industrial robot system via a traditional teach pendant and a DT leveraging a VR user interface. The study group involves forty subjects including experts in robotics and VR as well as non-experts. An analysis of the data gathered in both the real and the virtual use case scenario is provided. The collected information includes time for performing a task with an industrial robot, stress level evaluation, physical and mental effort, and the human subjects' perceptions of the physical and simulated robots. Additionally, operator gazes were tracked in the VR environment. In this study, VR interfaces in the DT representation are exploited to gather user centered metrics and validate efficiency and safety standards for modern collaborative industrial systems in I5.0. The goal is to evaluate how the operators perceive and respond to the virtual robot and user interface while interacting with them and detect if any degradation of user experience and task efficiency exists compared to the real robot interfaces. Results demonstrate that the use of DT VR interfaces is comparable to traditional tech pendants for the given task and might be a valuable substitute of physical interfaces. Despite improving the overall task performance and considering the higher stress levels detected while using the DT VR interface, further studies are necessary to provide a clearer validation of both interfaces and user impact assessment methods.

Keywords: digital twin; human–robot interaction; industrial robotics; virtual reality

Citation: Kuts, V.; Marvel, J.A.; Aksu, M.; Pizzagalli, S.L.; Sarkans, M.; Bondarenko, Y.; Otto, T. Digital Twin as Industrial Robots Manipulation Validation Tool. *Robotics* **2022**, *11*, 113. https://doi.org/10.3390/robotics11050113

Academic Editors: Carlo Alberto Avizzano and Kensuke Harada

Received: 27 July 2022
Accepted: 13 October 2022
Published: 18 October 2022

1. Introduction

There is a growing body of literature recognizing the importance of digital twin (DT) in numerous research fields. An increase in the number of publications involving industrial human–robot interaction (HRI) and human–robot collaboration (HRC), in particular, demonstrates the focus of DT research in which virtual implementations of physical robot cells enable safe and efficient tools for system evaluation, training, and offline programming [1]. Despite the growing prevalence of DT in such applications, however, there is relatively little known about the human factors that drive and impact DTs of manufacturing systems [1]. Significant efforts continue to move toward human-centric design and implementation. Villani et al. [2], for example, present an extensive overview on HRC in industrial settings focusing on the main topics of safety, applications, and intuitive human–machine interfaces (HMI). That study presents alternative solutions to the traditional interfaces (i.e., based on keyboards and mouse, or teach pendants), namely walk-through programming, teach-by-demonstration methods, multimodal natural user interfaces (NUI; e.g., vision-based gesture recognition and audible speech recognition), and augmented, virtual or tangible user interfaces (TUI). That survey further points out the advantages of these innovative approaches in terms of reduction of time and costs related to the robot programming task and safety assessment, while also highlighting the importance of evaluating human factors such as stress, workload, and mental safety.

A number of practical reasons for pursuing DT solutions of manufacturing processes exist; and the number of challenges facing such implementations are plentiful. Technology transfer, retrofitting legacy robots, and adopting novel digital technologies in the historically manual and analog systems leveraged by small- and medium-sized enterprises are among the principal challenges [2]. Documenting the approaches and impacts of introducing DT in existing workcells is, therefore, expected to be both illuminating and beneficial in future iterations.

This study presents an application of DT in industrial robotic applications with a specific focus on the human factors that drive utility and adoption of DT technologies. A design of experiments is proposed to capture both quantitative and qualitative data regarding operator use and preferences of physical and virtual interfaces. A study involving forty ($N = 40$) volunteers through Tallinn University of Technology is leveraged to evaluate and verify the test methodology and initiate a validation of the specific tools. The central hypothesis of the work is that the DT with enabled immersive technologies user interface can be adopted as a task/safety standards validation tool for Industrial robotic applications which involve human–robot interactions (HRI). The study aims at detecting if any degradation of user experience and task efficiency exists when using immersive user interfaces in comparison to real robot tech pendant. What we aim for is to introduce user impact evaluation within the DT VR interface prior to the actual adoption of the technology in real world use cases and at the same time find which are the most appropriate metrics to detect the efficiency and effectiveness of the proposed tools in respect to HRC tasks.

1.1. Related Studies

Several publications illustrate both the state of practice and emerging advancements in the field of DT augmented and virtual reality (AR/VR) interfaces for robot programming, training, and safety assessment. Nevertheless, not many works propose a standardization of evaluation methods for human factors and a ground base comparison of interaction efficiency between real and virtual environments able to define a set of metrics relevant and applicable to non use case specific scenarios. The horizon of studies, methods and tools in this respect is wide and heterogeneous. The following subsection attempts organizing them based on specific use cases and applications relevant for this study.

1.1.1. Input Modality Evaluation

The topic of NUI in HRI for industrial and service robots is discussed by Berg and Lu [3]. Their review mentions control interfaces based on gesture and speech recognition in combination with virtual and augmented reality, portable devices, or eye-tracking, highlighting the importance of a multimodal approach to HRC. The study by Krupke et al. [4] presents an experimental setup comparing mixed reality (MR) robot interaction and control based on heading position and direct selection, with speech input for task and action commitment. A virtual robot arm is used in a pick-and-place task, and is synchronized and superimposed over the display of a physical robot, allowing for movement preview in MR and facilitating the robot programming task and procedure. Experimental results confirm that heading-based selection of controls to be faster, more precise, and less demanding on the user. Their tests assess operator performance through commonly-used questionnaires, namely the National Aeronautics and Space Administration's Task Load Index (NASA-TLX, [5]), the System Usability Scale (SUS, [6]), the AttrakDiff Usability

Questionnaire [7], and objective metrics like completion time and accuracy. The study by Whitney et al. [8] compares 2D and DT based immersive VR interfaces aimed at robot control in a simple object stacking task. The evaluation of different control interfaces includes direct manipulation, keyboard and monitor, hand position tracking with monitor visualization, hand position tracking in combination with virtual reality. User tests, together with NASA-TLX and SUS responses, indicate that the VR interface is more efficient, faster, with a lower workload and higher usability than the monitor and keyboard one. Hand position tracking is also an important key advantage in robot manipulation in combination with both monitor based and VR visualization methods. Direct manipulation proves to be the best type of interface for the given task overall. A VR DT interface for aircraft engines performance control is presented by Tadeja et al. [9]. Information and nominal performance maps are synchronized with both real and digital representations of engines, allowing for real-time visualization and manipulation in the immersive environment. Several VR-based interfaces and interaction methods, such as pinch based hand manipulation and gaze tracking, are tested by a limited number of users in the performance of a specific engine inspection and control task using the proposed application. An extensive set of metrics and questionnaires are employed in this study to detect user health, workload and reactions to the virtual scenario, such as sickness and flow. Among the others the author mention the Simulation Sickness Questionnaire [10], Flow Short Scale [11], NASA-TLX and Igroup Presence Questionnaire [12]. The study by Laaki et al. [13] makes use of a virtual reality DT environment for the remote teleoperation of a Universal Robot UR3 robotic arm to simulate a remote surgery scenario. The study focuses on security, reliability of the connection over a mobile network and usability of the system. The importance of Quality of Experience (QoE) and the assessment of human factors, such as sense of presence, visual fatigue, cyber-sickness, and system acceptance in VR based teleoperation tasks is discussed by Concannon et al. [14]. The study presents a framework for QoE assessment by employing a DT simulation in VR and tries to establish the impact of network delay by using implicit and objective metrics. User physiological data such as heart rate, electrodermal activity, eye-tracking focus of attention and environment interaction variables are used as a base for user experience assessment.

1.1.2. Human Robot Collaboration and Work Cell Optimization

The literature is rife with examples of evaluations to demonstrate and assess DT in HRC applications. Matsas et al. [15] describe a VR HRC environment for the performance of complex tasks in a collaborative industrial use case. The VR scenario is enriched by audio-visual cues, cognitive aids, and interaction metaphors. The reported evaluation of the system gives positive results in term of acceptance. In particular, the users appreciated the system aids and cues, particularly when turning into potential danger and collisions warnings and alerts. Despite proposing several user experience evaluation metrics the study fails in providing a standardized assessment of user perception and experience in the system by utilizing a custom made questionnaire. Similarly, Oyekan et al. [16] explore the effectiveness of VR in developing HRC strategies. An experimental DT is employed in the evaluation of human reaction to unexpected robot movements while carrying out a human-fed, pick-up-and-transfer task. A variety of different metrics are considered, including head acceleration, head and neck energy metrics, direction and angle of human reaction and Head Injury Criteria-based force related danger of the robot movement. The study suggests that VR DT can help determining and understanding human reactions to robot movements facilitating the definition of HRC strategies in a safe and controlled environment. In this study, a custom questionnaire is employed to capture the users reactions to robot behaviours. Complimentary to this, DT is also leveraged as a tool for HRC task design and wok cell design optimization. An architecture for DT MR environment aimed training and based on a modular experimental collaborative robot assembly plant is described by Sievers et al. [17]. The study by Yap et al. [18] presents a VR projection-based environment for robot control and programming. Taking into account ergonomics parameters for each worker, the Virtual Wall hardware architecture includes head and hand-manipulator tracking, polarized glasses and filters, active stereo glasses with a non-polarized projection screen. The system aims at the design and optimization of the cell layout and, in a second phase, at the use of the validated setup in the robot programming task. Once validated the robot paths are transferred to the real robot for testing. The study by Malik et al. [19] describes a framework for a VR-based HRC process design. A DT of the robotic cell and a human avatar are employed for collision analysis, reach, vision and placement tests of the robotic cell modules and components. A virtual interface allows the user to interact with the environment and the robot end effector. The validated robot positions are saved for later use in the real world robotic cell and synchronized with the physical robot. The study in Pérez et al. [20] validates the layout design of a multi-robot industrial cell in VR making use of existing DT cell components. The experiments described in the study by Peruzzini et al. [21], attempt a holistic description and modeling of an operator involved in an industrial vehicle assembly task. The goal is to improve and optimize the assembly workspace and find corrective actions for possible emerging issues. This study constitutes an interesting example of assessment of the impact of the adopted technologies and evaluation methods on the user, and the validation of the design approach for the specific use case, being as well applicable to HRI scenarios. A large number of subjective and objective metrics are collected and analyzed during experimental sessions in the real scenario, in VR and in mixed reality. Bio sensors, motion capture, video recordings and eye tracking are involved in collecting information over heart rate, breath rate, temperature, eye gaze and pupil dilation, body position and movements. Both physiological and psychological response are employed in the assessment of mental workload, comfort, ergonomics parameters, posture, visibility, and occlusion. The occupational repetitive actions (OCRA, [22]) Index and RULA score are used together with heuristic analysis in the assessment of ergonomics of the workstation while NASA-TLX questionnaire is employed for subjective workload evaluation.

1.1.3. Ergonomics and Safety Evaluation

From a more safety and user-centric perspective, Harvard et al. [23] present a simulation and communication architecture intended to design and evaluate assembly lines, manufacturing processes, and workstations. The system employs a DT VR interface allowing for efficient and safe configuration and validation test of the workstation setup and ergonomics. The user based evaluation is, in this case, achieved by adopting a Perception Neuron Pro sensor-embedded suit, for body posture and skeleton detection, and leveraging the Rapid Upper Limb Assessment (RULA, [24]) ergonomics scoring tool. Likewise, an experimental comparison of robot collision prediction and control via direct supervision, monitor and mouse interface and a mixed reality system is presented by Rosen et al. in [25]. Experimental results show that the MR interface is significantly more efficient, direct, and easier than monitor visualization and control while not being significantly less efficient and usable than the direct supervision of the robot. The study makes use of the NASA-TLX and SUS to determine user workload and assess system usability for the three interaction methods. With respect to user training and safety assessment in a HRC industrial setting, Moniri et al. [26] propose a remote collaborative setup supported by eye-tracking and virtual reality. The system is meant for online remote tutoring, training, and assistance. The experimental setup encompasses two synchronized workstations, a real and a virtual one. The system can track the position of the objects involved in the task, the robot manipulator orientation, user head position, and eye movements. Focus of attention information for each user can be visualized by the remote assistant during the pick-and-move task. Object collision avoidance is discussed by Wassermann et al. [27]. The monitor-based augmented reality application allows for the visualization of bounding boxes around real objects involved in a pick-and-place task. The system can detect and visualize the collision of the virtual robot with the bounding boxes by changing their color in real-time. Several robot safety behaviors are tested by Vosniakos et al. [28] making use of the virtual environment presented by Matsas et al. [15]. The study explores the effectiveness of HRC collision avoidance methods such as speed reduction and move back strategy. Similarly, the work by Maragkos et al. [29] aims at developing a collision avoidance system based on slow down strategy and alternative movement path on a traditional industrial robot for safe HRC and programming. The system implements a VR DT of the real robot where the virtual space is mapped and subdivided in 3D regions which are subsequently checked for the presence of the human avatar. The experimental study presented by Manou et al. [30] adopts a virtual reality robotic cell twin for the assessment of collision detection and robot movement paths generated by manually operating a sensor-enabled teaching tool in a lead-through programming method session.

1.1.4. Robot Programming

Similarly, DT is often leveraged for programming robotic tasks. A virtual environment for training simulation and programming in proposed by Pérez et al. [31]. The operator can interact with the robot through a virtual interface and assess the efficiency and safety of the proposed robot trajectories. The system stores paths and trajectory information from the virtual robot for further data analysis, training, and real robot programming. A custom questionnaire is employed to assess users experience of the system. A mixed reality robot programming interface making use of HoloLens is presented by Ostanin and Klimchik [32]. A virtual robot is programmed by a set of AR interfaces, manipulators and tools controlled by gesture inputs (tap, tap and hold). The system is tested for object avoidance and the creation of linear, circular and rectangular task programming paths. The AR interfaces allow the operator to modify (erase and scale) the proposed paths and directly control the robot's joints movements. The study in Nathaneal et al. [33], compares performance metrics over different user groups programming a robot by means of a traditional teach pendant, a non-immersive virtual environment and a virtual-augmented system. The user performance evaluation focuses on timing, the number of coordinate axis changes, and optimal piece positioning with the end effector. Several signals and alerts are implemented in the virtual environment to facilitate the robot programming task. Experimental results demonstrate that programming performance and time would benefit from the augmented cues and signals implemented in the system. The study suggests that the skills developed in the VR environment are transferred in the real case scenario by facilitating learning of traditional interfaces and robot manipulation. A lead-through offline programming approach based on augmented interfaces and a handheld pointer is presented by Ong et al. [34]. The pointer is directly operated by the user to create and modify paths related to different manufacturing tasks. Graphical cues real-time information about manipulability and reachability for each proposed path. The application is tested on a group of users confirming the usability of the system and a reduced amount of programming time compared to traditional methods. User experience and system usability are assessed based on a custom questionnaire and by comparing quantitative data collected during different experimental sessions. A VR based system for robot programming in a collaborative scenario is presented by Burghardt et al. [35]. Unfortunately, the study does not provide quantitative data analysis on the comparison of traditional programming methods and the proposed application.

Several publications propose an assessment and evaluation of the effectiveness of interaction methods and hardware for DT interfaces, but only a few try to compare traditional robot programming methods and immersive VR DT solutions. In this sense, it is important to determine whether there is a degradation in the use of DT VR interfaces compared to real robot teach pendants and establish if the former can be reliable and efficient substitute for HRI. Moreover, the evaluation of human factors and the impact of the system on users' interaction with VR interfaces are not frequent and, in most of the cited cases, use case specific. Several aspects are determining the efficiency and effectiveness of VR interfaces for HRC including the acceptance of the system, usability, users' stress level, and workload. By analyzing specific metrics in a real HRI and in a DT VR scenario, this study aims at the comparison of performances in robot programming using traditional and VR interfaces and the evaluation of their impact on the operator. We believe this type of assessment should be performed prior to use case specific applications, addressing the need for metrics and validation methods which would acknowledge the centrality of the operator in the DT loop as part of the VR interface and at the same time address the effectiveness of a DT system for robot control. To do so it is necessary to compare the control efficiency and user interactions with the interface in the virtual environment compared to the physical workspace and evaluate what is the impact of both on the operators. The hypothesis is that the DT VR tools can be as efficient as their real counterpart with minor impact on the user health, stress levels and performance indicators.

2. Methods

The experiments presented in this study consist of both physical and virtual tests in identical work cells as showed in Figure 1. Both the physical and virtual robot configurations were used to complete identical material handling tasks. The tasks consisted of moving three cubes–located in different parts of the workspace to a predefined target region. Each cube has a predefined starting position and must be picked up and moved to the target region in a specified order. The robot is teleoperated by the human subject using interfaces specific to the operating environment: experiments using the real robot were performed using the teach pendant provided by the robot's manufacturers (see Figure 1), and experiments using the virtual robot leveraged a custom user interface displayed in the simulated environment (see [36], and Section 2.3.1 for the description of this interface.

Figure 1. The design of experiments utilizes both physical and virtual representations of the work cell. The physical trials (**left**) involve the operator using the robot's teach pendant to manipulate the end-of-arm tooling. The virtual trials use a VR interface (**right**) for commanding robot motions.

2.1. Human-Robot Interaction Metrology

Several different metrics and test methods for assessing and assuring HRI technology performance are detailed in the literature (e.g., [37]). This broad spectrum of metrology tools can make selecting appropriate test methods a significant challenge. Given this study's focus on the use of VR for HRI, a collection of metrics that capture interface utility and operator reactions to the interface controls are warranted.

To capture the human operator' interactions with the interface, a combination of quantitative and qualitative metrics (both objective and subjective) are selected. The objective measures capture the nuances of user interaction with the interface that may not be registered or recollected by the operator in a post-test questionnaire assessment. In contrast, the subjective measures capture the in situ effects of interface interaction and manipulation for a specific individual at a particular time during or after the test. Given that external and personal factors (e.g., the effects of weather, diet, and recent events on the individual's temperament and focus) are often influencing subjective results, it is generally advisable to consider them as anecdotal samples of a larger range of random responses rather than an absolute constant. Therefore, the trends of these responses are more indicative than the responses themselves.

The objectively quantifiable measures selected are intended to capture how the operators actually use the interface. In particular, the measures identify three important factors: (1) how much time was required to complete the task, (2) where the operator's focus is predominantly drawn to throughout the task, and (3) how much time is spent focusing on these elements. To achieve this, the following factors are captured and reported in this report:

- the average time to complete the task working with the physical robot versus working in the virtual environment;
- the average total duration of the experiments using the VR interface including both the time to complete the task and the time to adjust to the virtual environment;
- the average duration using or focusing on the different virtual interface commands (e.g., adjusting joint speeds/positions or changing operational modes);
- the average total time spent looking at the elements of the VR interface, the time spent looking at the virtual robot, and the time spent doing something else (e.g., doing some other task work not directly related to the robot).

Only the first of these metrics is captured for both the physical and virtual robot work cells. No reliable and repeatable system currently exists for tracking operator attention or gaze focus for real-world interfaces. In contrast, operator eye motions can be reliably tracked in the virtual environment using VR headset-mounted eye trackers. As such, the sampling of objectively quantitative measures can be built directly into the interface itself.

Two subjectively quantifiable measures were selected to capture the users' experiences of using the interfaces. It has been seen that exposure to and experience with robots has an impact on the users' responses to robots. For example, if the operator has plenty of exposure to robots (e.g., as an influence of popular culture representations such as in movies) but little practical experience working with them, these users may over-estimate the robots' intelligence and capabilities. Similarly, users who have both little exposure and little experience with robots may express fear or excitement working with the robots. To this end, the following metrics were captured:

- the demographics of the human operators, including age, gender, and nationality;
- the operators' previous experience with working with both robots.

Finally, two popular, subjectively qualitative survey tools are selected to probe the operators' reactions and opinions of working with the robots in both the physical and virtual work spaces. These surveys are intended to capture the operators' perspectives on the difficulty of using a given interface, and the operators' perspectives on how they felt around the robots:

- the NASA-TLX captures the operator's mental and physical effort required to complete a task;
- the Godspeed questionnaire [38] records the operators' perspectives of the anthropomorphism, animacy (i.e., how lifelike something appears), likeability, perceived intelligence, and safety of the real and virtual robot systems.

It is worth noting that additional subjective software quality metrics and test methods have been standardized in ISO 25010 ("Systems and software engineering-Systems and software quality requirements and evaluation (SQuaRE)-System and software quality models," [39]). However, these test methods and metrics capture only the user's perspective on the quality of the software (i.e., the interface), and does not reflect the user's experience using the software for interacting with robots.

2.2. Data Collection: Integrating Metrics

Each element that can be manipulated (i.e., buttons and sliders) on the VR interface canvas is assigned a unique tracking identifier. Each UI function controlled by the operator has an attached script that identifies with which feature the user is currently interacting, as well as the type of ongoing interaction (e.g., button press/release). The operator's direction of gaze is estimated by casting a virtual line, originating from the normal surface of the head-mounted display (HMD), extending outward to the virtual world. During the experimental session, such setup allows to record any user's interactions with the UI in a timeline, which is then saved to a JavaScript Object Notation (JSON) file detailing all events on a per-element basis.

A script simulating network instability and lag was also introduced for the virtual robot. This script injects arbitrary disturbances into the visual representation of the DT, with the goal of affecting the operators' behavior. The generated experimental session file can be later used to analyze the performance of each operator. For example, an activity heatmap can be generated on top of the UI image to visualize the length of time spent in each UI section, or the amount of interactions with each unique element.

Both the NASA-TLX and the Godspeed metrics are post-task surveys, so do not provide real-time data collection. The inputs from these surveys are then assessed in an effort to map the operators' responses to their respective interactions with the interface.

2.3. Technical Implementation

To evaluate the differences between the experience of the operators when working with the real robot system and its digital counterpart, an existing DT system was augmented with tools to track behavior metrics of the users, as described in Section 2.2. Several tracking metrics were used when collecting data during the experimental sessions. These include timing, and the operators' attention and stress levels. A detailed description of the experimental configuration is described presently.

2.3.1. The Digital Twin System

The DT system used as a basis for this experiment was developed at the TalTech Industrial Virtual and Augmented Reality Laboratory (IVAR) during the previous research on the relevant topic (see [36,40,41]). The system was developed using the Unity game engine, and contains a digital model of an industrial robot that can be manipulated using the accompanying UI located in virtual environment (see Figure 2).

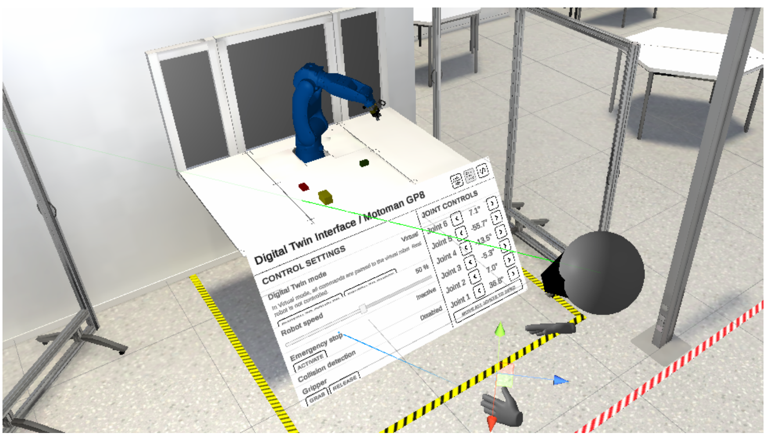

Figure 2. The user interface of the industrial robot cell DT. The interface is a large panel located in front of the robot, and is manipulated using a hand-held pointing device.

The model of the robot is retrieved from the manufacturer's website, and went through an optimization procedure including rigging and the creation of rotation axes, mesh simplification, and scale correction. These last steps are necessary to ensure that the imported mesh would be identical to the objects in the real scenario. Hierarchical structures between rig pivots (robot axes) is maintained and based on the original robot technical drawings. Precision of joint controls in the virtual environment is identical to the real robot precision with accuracy of 0.001 degrees, and joint limits are set to be identical to the real counterpart. The speed of the synchronized real robot is proportional to the one set by default in the DT VR environment. Although the physical robot has built-in collision detection (which was triggered several times throughout the experiments), no collision avoidance scripts were used in the DT counterpart.

The DT can be operated in "coupled" and "virtual" modes. In the coupled mode, all commands are duplicated and sent to the physical robot over the local network, effectively keeping the virtual robot synchronized with its real-world counterpart. In the virtual mode, the network link between the DT and the real robot is disconnected. All actions happen inside the simulation only. Apart from these connection modes, the UI provides two control mechanisms for commanding robot motions. The user can either directly tele-operate the robot arm by adjusting individual joint positions, or create a multi-step, joint-space program to be stored and executed later.

For this experiment, the virtual mode with direct control was chosen for two reasons. First, the DT operated in virtual mode is not bound to the physical speed and safety limits set in the real robot system, which allows an unbiased assessment of the possible performance benefit of DT solutions. Second, controlling the virtual robot directly is similar to controlling the real robot with its included teach pendant, and does not introduce additional complexity in the form of creating and executing program. This was particularly important, as the collection of human operators participating in the experiment represented a wide spectrum of prior experiences and expertise with robot systems. Using direct control eliminates unnecessary complexity when preparing for the experimental session.

2.3.2. Attention Tracking System

An attention tracking system was developed to record the behavior of the human subjects during the experiment. This system allows "tagged" objects in the virtual environment as attention targets, and produces reports on how long the user's attention was directed to the specific object and at what specific moments in time. The system's principal architecture diagram is shown in Figure 3.

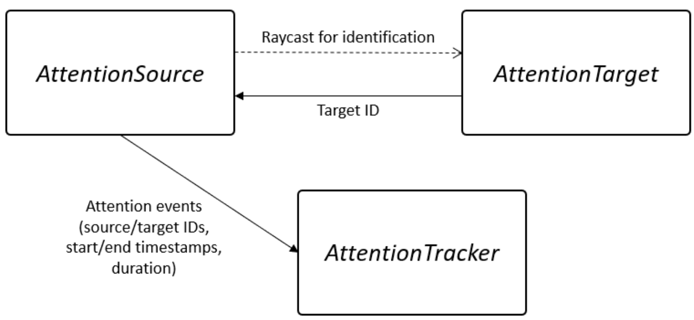

Figure 3. Architecture of the attention tracking system.

The attention tracking system consists of several components, each of which corresponds to a single script written in C#. The purposes of these components are as follows:

- *AttentionTarget*—a script which marks an object as a target for the attention tracking system. It is a Unity Component script, which means it can be attached to any 3D object in the virtual environment. *AttentionTarget* must have a trigger volume attached to it to be detected by *AttentionSources*.

- *AttentionSource*—a script which is responsible for detecting *AttentionTargets*. *AttentionSource* uses raycasting to detect trigger volumes with associated *AttentionTargets* in the virtual environment. If the "line of sight"–originating from the normal surface of the HMD–check encounters an *AttentionTarget*, a new attention event is triggered for that object. Once this object is no longer along the line of sight for some n number of computational cycles, the attention tracking event is considered finished and its duration and timestamps (in milliseconds) for the beginning and end of the event are written into the session file for the *AttentionTracker*.The precision of events duration is equal to the simulation's clock cyclek length. Here, the environment used in the experiment is executed at 120 Hz, which yields a maximum precision of 8.3 ms. *AttentionSource* is thus leverages head tracking as an approximation of eye tracking and attention monitoring.
- *AttentionTracker*—a core script which provides methods to start and stop the recording of the attention tracking session, register attention events, and export recorded data in JSON format for later retrieval and analysis.

The DT interface is segmented into three primary zones (see Figure 4), two of which are used for controlling the robot's actions. The interface's header draws the operator's attention to the robot being controlled using the DT interface. The general controls section is used to adjust system settings, including robot speeds, activating/deactivating the robot, and actuating the gripper. The joint controls section is used to adjust the orientations of the individual robot joints, starting at the tool flange (Joint 6) and moving down the kinematic chain to the base (Joint 1).

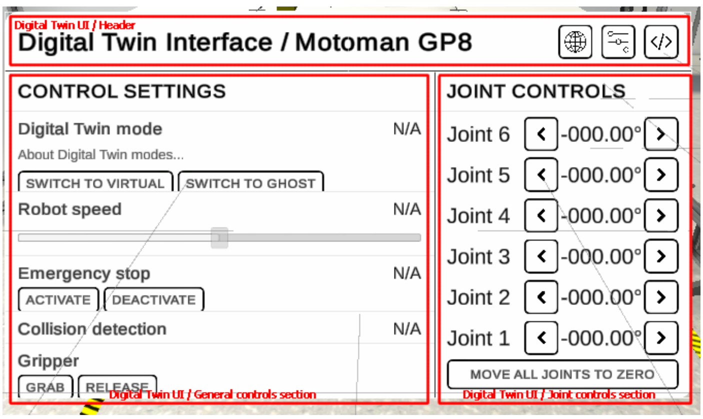

Figure 4. The three primary zones of the DT UI are the header (**top**), the general controls (**left**), and the joint controls (**right**).

For this experiment, a Vive Pro Eye VR headset with the Vive SRanipal software development kit was used to access eye tracking data in Unity engine. Eye tracking data is usually provided as a direction of the user's gaze, expressed in quaternions. A custom C# script was used to apply gaze direction data to objects with *AttentionSource* scripts attached. As a result, the operator's gaze could be used directly to register attention events. This approach is not limited to a specific HMD model, and can be replicated using other eye tracking systems. Eye tracking is used as a measurable proxy of operator attention within the virtual environment. When using the virtual interface, eye tracking can also be leveraged as an objective measure for estimating ease-of-use of the interface (e.g., how much time is spent scanning the interface for the appropriate functions for commanding robot motions), optimizing cues for attracting operator attention (e.g., are there any visual elements that distract the operator from their work?), and can benefit future iterations of the interface design (e.g., by clustering common functions as determined by frequency analysis, or by pattern identification of common gaze shifts).

2.3.3. Stress Level Approximation

To assess stress levels of human operators in both the real and virtual environments, the operators' heart rates were sampled during the experimental sessions using a heart rate monitor. The data was recorded for each person and then analyzed according to the flow of the experiment, identifying the operator's reaction to different experiment stages (introduction, explaining controls, and executing the task). All heart rate readings were supplied with timestamps to help categorize the data. These values can thus be compared with the resulting NASA-TLX results to provide quantitative estimates of physical and mental loads.

3. Experimental Protocol

As described in Section 2, an industrial robot motion control routine was chosen as a candidate use case for assessing the human operators' interactions with both the robot handheld joystick and the DT VR interface. All trials are based on a simple material handling task in which the human operators remotely control a robot to pick and place a sequence of blocks within the robot's work volume. Both the physical and virtual interfaces have similar capabilities in that they allow the operator to move the robot's joints, and actuate the gripper to pick up or release the blocks.

3.1. Robot Control Task

Identical tasks are used in both experimental conditions: users must control the robot to move objects, in order, from their initial positions to a target region. An operator is asked to sort the three objects by using an interface to control the robot to move to the objects, pick up the objects, and re-position the objects to a target location on the table in front of the robot (see Figure 5). Wooden cubes with different sizes were chosen as representative objects for the task. The blocks are both color-coded and labeled such that the order of the blocks is known. The operators were given minimal training on using the physical and virtual interfaces such that they could become acquainted with the controls, but not necessarily adept at using them. Following this initial training, the operators were then instructed to perform the material handling task using the interfaces specified for the operating environment. During the trials, the operators were not permitted to approach the robot. Each operator was instructed to complete the task using the physical robot, and then complete the task again using the virtual robot. After each task completion, the operator would complete the NASA-TLX and the Godspeed questionnaires. The three cubes started in the same initial poses for all trials.

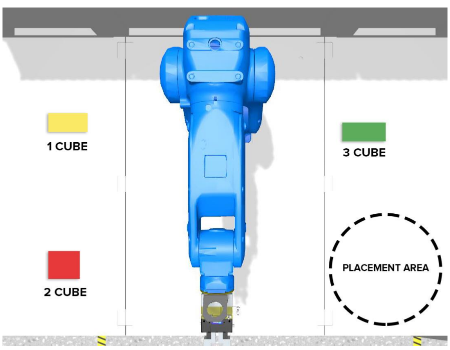

Figure 5. Configuration of the the block manipulation task experiment.

For each block, the task consists of moving the robot's gripper such that it could grasp the block. This often necessitated some trial-and-error, which could increase the completion time, but was also expected to result in net performance improvements as a function of time. Precision was necessary for grasping and lifting the blocks, but placement of the blocks did not require as much accuracy or repeatability. All robot motions after acquiring the blocks were performed in free space without obstacles, allowing the researchers to capture best-case timing such that only the operator's ability to use the interfaces impacts the task performance. During the trials, the operator's actions and performance were recorded to assess task performance and interface utility.

3.2. Participants

A total of forty-seven subjects volunteered for the trials, but seven volunteers did not complete the trials and are therefore excluded from reporting (therefore serial number of subjects throughout the paper is not continuous). The remaining volunteers (N = 40) were then divided into two equal groups, which differed in terms of the order in which the subjects used the different interfaces. One group evaluated the physical teach pendant interface before the DT interface, while the other evaluated the DT interface first.

For the first volunteer group (Group A, physical interface first, $N_A = 20$), sixteen (16) participants identified as male, and four (4) identified as female. The sample included backgrounds from engineering, business administration, and environmental engineering bachelor and master students; engineering disciplines lecturers and researchers. The number of selected subjects was limited due to the COVID-19 quarantine period and consequent restrictions in human gatherings. The age of subjects ranged between 20 and 53 with an average age of 29.4 years. The subjects' countries of origin were divided as follows: seven (7) from Estonia, five (5) from Ukraine, two (2) from Iran, two (2) from Turkey, one (1) from Bhutan, one (1) from Georgia, one (1) from Nigeria, and one (1) from Pakistan. Users were asked to evaluate their skills in robot programming on a scale from 1 to 10 (1 being no experience, and 10 being an expert in robotics). The average response for this group was 3.9, showing low self-assessment grade in the field of related research.

For the second volunteer group (Group B, DT interface first, $N_B = 20$), seventeen (17) identified as male, and three (3) identified as female, and had an average age of 29.9 years old (with a range from 22 to 53 years old). Volunteers had different backgrounds being mainly students and researchers from different departments of Tallinn University of Technology, and one professor. The sample included eight (8) people of Estonian nationality, six (6) from Italy, one (1) from India, one (1) from Slovakia, one (1) from Turkey, one (1) from Japan, one (1) from Ukraine, and one (1) from Ecuador. The average self assessment value of expertise in robot programming scored, this time, a value of 4.2, which represents, again, quite a low expertise estimation.

Volunteers were instructed to complete the robotic object handling task over three trials for both interfaces, resulting in six trials in total per person.

4. Results

4.1. Task Timing

Table 1 shows a comparison of average completion times for both the physical and virtual interfaces for volunteers in Group A. By contrasting the average time spent by volunteers in the real-world trials versus the DT interface, it is seen that the use of the physical teach pendant generally resulted in significantly longer times to complete the robotic material handling task. This trend is observed for the manipulation of all three cubes.

Table 1. Average task completion time, in seconds, for Group A (physical interface first).

Cube Re-Positioning Task	Average Physical Machine Duration	Average VR Process Duration
Cube #1	225	144
Cube #2	210	115
Cube #3	184	105

Table 2 shows a comparison of average completion times for volunteers in Group B. It is seen that the resulting trends in average completion times are comparable with those in group B, with times to perform the task using the virtual interface being less than the times using the physical interface. It is also observed that Group B demonstrated better task execution performance overall. This latter observation could imply that VR-based experiment introduction and testing is more beneficial from the perspective of preparing the users to work real machinery. However, it may also be a result of the slightly higher average self-reporting robot expertise score than Group A.

Table 2. Average task completion time, in seconds, for Group B (Virtual interface first).

Cube Re-Positioning Task	Average VR Process Duration	Average Physical Machine Duration
Cube #1	114	178
Cube #2	92	130
Cube #3	61	159

To test this, two consecutive trials of volunteers performing the object handling task using the VR interface are evaluated for a subset of the volunteers. The volunteers are identified by their self-reported expertise in robotics. As seen in Table 3, which shows the task completion time using the VR interface for two consecutive trials, there does not appear to be a strong correlation between self-reported expertise in robotics and initial task performance. For example, one volunteer who self-reported their expertise level as “2” performed consistently better than a volunteer with a self-reported expertise of “9.” Some other factor (possibly experience with other machinery, video games, or similar systems) must be contributing to this discrepancy. Furthermore, the field of robotics-aerial, ground, industrial etc, were not asked as well as expertise with Virtual Reality applications, which could lead to faster learning curve towards the immersive experience of the experiment. Such information, however, was not captured in the initial surveys, and will be a subject of future study.

Table 3. Sub-task completion times (in seconds) of subjects with different self-reported experience in robotics using the VR interface.

		Trial 1		
Expertise in robotics (1–10)	2	9	1	2
1 cube end	360	120	180	120
2 cube end	240	120	120	90
3 cube end	60	180	120	90
		Trial 2		
1 cube end	300	120	60	60
2 cube end	240	60	60	45
3 cube end	60	45	60	45

Regardless, it is evident that, as the volunteers use the interface more, the times to complete the task are, generally speaking, monotonically decreasing. Similar trend lines are also seen while using the physical interface, as shown in Figure 6, it is also decreasing with the same average tempo. As is seen on the figure, on average the virtual trials took less time to complete than the physical trials. This may be attributed to some experience gained during the physical trials being applied to the virtual interface and person mental readiness for the future step. Moreover, virtual environment might seems more simple for the users due to it is similarity to the computer game rather to standing next to the physical machinery.

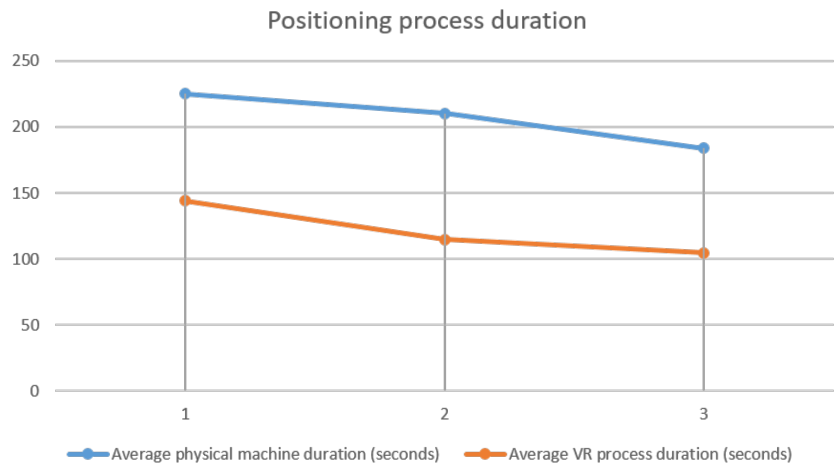

Figure 6. Average positioning time comparison per cube for both physical and virtual interfaces.

4.2. Subjective Survey Responses

Following the trials, volunteers were asked to complete the Godspeed Survey, which captures individual perceptions and reactions to robots after interacting with them, and the NASA-TLX, which is used to measure the physical and mental effort required to complete the task.

The averaged results from Group A's Godspeed surveys are given in Table 4. These results reveal no significant differences between the physical and virtual systems in terms of operator responses. Controlling the physical robot resulted in higher average results in terms of evaluating anthropomorphism, while the virtual environment was considered more interactive than the real setup. Likewise, the scores for perceived intelligencewere higher for the VR environment. Although the virtual environment was perceived as creating more anxiety than the real robot cell, the volunteers' perception of safety of the two systems shows no significant differences between the two scenarios. To test for any potential impact on the order of exposure (real versus virtual), the Godspeed survey for Group A was compared with that of Group B. Results for Group B, shown in Table 5, show that while the values are slightly different between the two groups, the overall trends do not differ significantly. This implies that the order of interface experiments (physical or virtual) does not affect overall perception of the robot systems.

Table 4. Godspeed survey results comparison, real versus virtual robots, for Group A.

Anthropomorphism-Scale 1–5	**Physical**	**Virtual**
Fake-Natural	4	2.65
Machine-like-Human-like	2.6	2.5
Unconscious-Conscious	2.45	2.3
Artificial-Lifelike	2.55	2.15
Moving Rigidly-Moving Elegantly	3.25	2.75
Animacy-Scale 1–5	**Physical**	**Virtual**
Dead-Alive	2.5	2.4
Stagnant-Lively	2.85	2.85
Mechanical-Organic	2.3	2.25
Artificial-Lifelike	2.3	1.9
Inert-Interactive	2.75	3.35
Apathetic-Responsive	3.4	3.45
Likeability-Scale 1–5	**Physical**	**Virtual**
Dislike-Like	4	3,.5
Unfriendly-Friendly	3.55	3.3
Unkind-Kind	3.4	3.35
Unpleasant-Pleasant	3.6	3.55
Awful-Nice	3.9	3.55
Perceived Intelligence-Scale 1–5	**Physical**	**Virtual**
Incompetent-Competent	3.2	3.3
Ignorant-Knowledgeable	2.95	3.45
Irresponsible-Responsible	3.45	3.5
Unintelligent-Intelligent	2.9	3.2
Foolish-Sensible	3.1	3.25
Perceived Safety-Scale 1–5	**Physical**	**Virtual**
Anxious-Relaxed	4.1	3.55
Agitated-Calm	3.85	3.9
Quiescent-Surprised	2.95	3.05

To contrast the results of the Godspeed questionnaire, Group A's NASA-TLX survey results, averaged and shown in Table 6, demonstrate that the use of the VR programming environment was considered more mentally demanding, and created a higher level of frustration and required more effort than the real environment. Performance evaluation of the physical trials was also slightly higher than the virtual trials. Group B's NASA-TLX results, Table 7, demonstrates a reversal in the perception of effort, with the physical system largely demanding more effort and resulting in higher frustration than the virtual system. This demonstrates a correlation between the order of trial evaluations and the perception of effort. Namely, that the interface the participants experienced first tended to be perceived as demanding less effort, but ultimately performed worse, than the second interface. This could infer a potential resistance to change, particularly when introducing new technologies in established processes. Additional experiments will be necessary to confirm this.

Table 5. Godspeed survey results comparison, real versus virtual robots, for Group B.

Anthropomorphism-Scale 1–5	**Virtual**	**Physical**
Fake-Natural	3.55	3.45
Machine-like-Human-like	1.95	1.9
Unconscious-Conscious	2.35	2.3
Artificial-Lifelike	2.45	2.2
Moving Rigidly-Moving Elegantly	3.05	3.45
Animacy-Scale 1–5	**Virtual**	**Physical**
Dead-Alive	2.9	2.45
Stagnant-Lively	3.2	3
Mechanical-Organic	2.25	1.85
Artificial-Lifelike	2.4	2.05
Inert-Interactive	4.1	3.35
Apathetic-Responsive	4.25	3.7
Likeability-Scale 1–5	**Virtual**	**Physical**
Dislike-Like	4.45	3.8
Unfriendly-Friendly	3.95	3.1
Unkind-Kind	3.7	3.1
Unpleasant-Pleasant	3.95	3.35
Awful-Nice	4.3	3.65
Perceived Intelligence-Scale 1–5	**Virtual**	**Physical**
Incompetent-Competent	3.65	3.2
Ignorant-Knowledgeable	3.45	3.05
Irresponsible-Responsible	3.3	3.15
Unintelligent-Intelligent	3.2	2.8
Foolish-Sensible	3.35	3.15
Perceived Safety-Scale 1–5	**Virtual**	**Physical**
Anxious-Relaxed	3.7	2.6
Agitated-Calm	3.8	2.95
Quiescent-Surprised	3.7	3

Table 6. Comparison of the average results from the post-task NASA-TLX surveys for Group A. Participant responses are given on a Likert scale of 0 to 10, representing the ranges shown in the left column.

Criteria	**Scale**	**Physical**	**Virtual**
Mental Demand	Low-High	4.55	5.5
Physical Demand	Low-High	2.925	2.575
Temporal Demand	Low-High	4.35	3.75
Performance	Good-Poor	3.125	3.5
Effort	Low-High	3.425	4.325
Frustration	Low-High	3.175	4.125

Table 7. Comparison of the average results from the post-task NASA-TLX surveys for Group B. Operator responses are given on a Likert scale of 0 to 10, representing the ranges shown in the left column.

Criteria	Scale	Virtual	Physical
Mental Demand	Low-High	4.35	6
Physical Demand	Low-High	2.25	4.65
Temporal Demand	Low-High	4.07	5.55
Performance	Good-Poor	3.85	4.55
Effort	Low-High	4.07	6.25
Frustration	Low-High	2.1	4.75

4.3. Eye Tracking

By comparing eye-tracking data collected during VR robot programming tests (Figure 7 and Table 8 for Group A, and in Figure 8 and Table 9 for Group B), it is clear that time spent looking at the UI controls is considerably higher (more than double) than time spent in looking at the virtual robot. For Group A, the regions containing the general controls, joint controls, and header are among the virtual UI targets with highest focus times. For Group B, the attention tracking system gave slightly different results. Figure 8 shows how the joint control section is still the area that was visualized for the longest time. In this case, though, timing relative to the general control section and header are much lower. In contrast, time spent looking at the physical industrial robot joint 6 is much higher than in the previous experiment.

Higher times for looking at robot joint 6 are quite understandable, as this joint is relative to the robot gripper and consequently the object to be picked and replaced. In contrast, there is no clear correlation between expertise with robots and time spent in looking at the controls or the robot in the DT. Given the comparatively short duration of the trials, drawing conclusions from the operator's extended use of different interfaces and the focus of their gaze is inconclusive. Future works could include the evaluation of user expertise with immersive VR technologies over longer periods of time.

While the raw numbers are interesting and telling in and of themselves from an individual participant's perspective, they do not succinctly capture the general performance of the operators during the experiment. By re-evaluating each time factor as a percentage of the total time spent using the interface, the data becomes normalized. When plotted as is shown in Figures 9 and 10 it becomes clear where the operators' attention was generally focused. Per Figure 9, a disproportionate amount of the operators' time was spent looking at the UI header, followed closely by the joint control panel. However, it is unlikely the operators' attention was focused this much on the header. As such, it can be surmised that the implementation of the eye tracker was somewhat flawed, with the most likely source of error being the assumption the operator's focus is determined exclusively by the positioning of their head. The more plausible hypothesis is that the operators' faces were pointed at the header (which is situated between the robot and the interface), and their eyes would move up or down to adjust focus on the joint controls and the robot.

A more precise implementation of eye/focus tracking would be to correlate the head position with the motions and interface usage of the pointing devices. For example, an extended period of time spent looking at the header while the pointing device is interacting with a button on the joint control panel can indicate focus on either the joint control panel, or on the robot's joint(s) being manipulated. If there is an extended period of activity (e.g., moving the pointing device or rapidly pressing the action button), one might assume the focus was on the joint control panel. Otherwise the virtual joint is the more likely target of attention.

Table 8. Eye tracking analysis of subjects use times (in seconds) for Group A, as reported as the digital twin UI, the virtual robot, and the operator's digitized body parts (Self; here, it is the operator's hand). The UI is segmented into the general controls, header (H), and joint controls (JC). The general controls are further segmented into emergency stop (ES), mode selection (MS), and speed selection (SS). The Robot is segmented per major component, specifically the base, joints 1–6 (J1–J6).

	Target													
	User Interface						**Robot**							**Self**
	General Controls													
Sub. #		**ES**	**MS**	**SS**	**H**	**JC**	**Base**	**J1**	**J2**	**J3**	**J4**	**J5**	**J6**	**Hand**
nr.3	38.49	15.28	13.49	5.77	21.57	117.46	15.77	13.72	25.98	5.48	50.79	60.5	99.12	0.13
nr.5	244.86	7.7	201.4	44.38	12.52	534.92	0.66	0.14	0.1	0	0	0	0	0
nr.6	22.72	2.15	13.62	7.56	112.42	28.82	39.08	53.49	50.04	2.38	6.46	2.39	4.24	0
nr.7	0.96	0	0.7	0	9.06	0.26	15.2	13.2	6.39	1.27	1.87	0.32	2.57	0
nr.8	16.19	0	6.61	0	51.49	1.8	5.34	2.08	2.04	0	0	0	6.25	0
nr.9	38.9	0	19.8	1.72	114	62.9	0.79	0.21	1.37	0.73	0.47	0.13	2.58	0
nr.10	29.5	0	22.97	5.11	208.51	41.03	112	160.19	194.15	39.03	71.7	26.82	29.46	3.84
nr.11	60.7	13.6	14	6.9	21.1	281	13.4	24.2	77.9	26	161	110	86.3	1.51
nr.15	34.94	0.48	27.36	7.89	71.44	90.32	2.37	2.23	4.1	0	0	0	0	1.84
nr.16	131	16.1	71.6	31.7	120	173	13.9	4.01	2.69	0.04	0.1	0.37	8.5	0.12
nr.17	22.02	0	13.88	0	177.68	22.6	2.36	0.69	3.83	1.26	1.79	1.29	2.6	0
nr.19	16.9	3.14	8.76	6.17	52.8	31.5	57.5	52.4	51.7	0	0.48	0.71	2.58	0.17
nr.20	15.72	0	13.35	2.07	51.94	26.27	4.72	7.17	14.05	3.56	16.57	7.74	8.08	0
nr.21	55.1	9.35	34.5	16.3	17.9	115	0.18	0.15	0.12	0	0	0	0	1.86
nr.22	1.92	0	0.11	0	9.16	0	0.92	9.57	34.42	31.08	11.64	2.6	2.57	0
nr.23	44.78	6.86	29.87	8.76	2.14	141.4	0.35	0.02	0	0	0	0	0	0
nr.24	35.12	0	12.34	1.27	199.63	26.83	21.24	4.36	4.24	0.06	0.64	0.36	1.1	0
nr.25	58.4	3.91	40.74	3.1	79.99	43.5	23.65	55.07	127.48	46.86	99.97	40.28	66.29	6.25
nr.26	26.73	9.09	12.81	9.97	49.97	60.19	4.11	0.96	0.73	0	0	0	1.48	0
nr.27	130.74	8.55	119.3	54.72	6.71	79	1.02	1.1	1.1	0	0	0	0	0

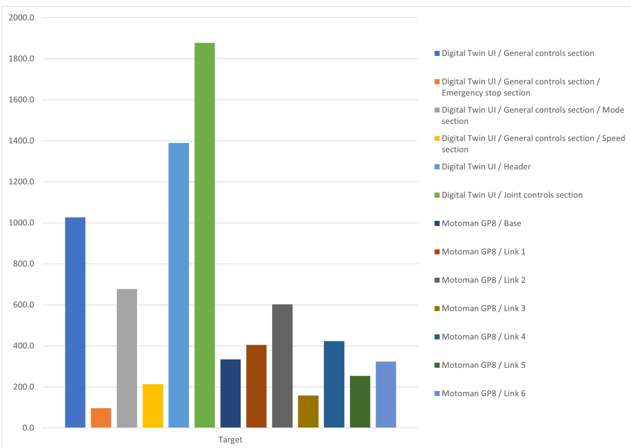

Figure 7. Total seconds spent viewing targets in VR for Group A

Table 9. Eye tracking analysis of subjects use times (in seconds) for Group B, as reported as the digital twin UI, the virtual robot, and the operator's digitized body parts (Self; here, it is the operator's hand). The UI is segmented into the general controls, header (H), and joint controls (JC). The general controls are further segmented into emergency stop (ES), mode selection (MS), and speed selection (SS). The Robot is segmented per major component, specifically the base, joints 1–6 (J1–J6).

Target														
	User Interface						Robot							Self
	General Controls													
Sub. #		**ES**	**MS**	**SS**	**H**	**JC**	**Base**	**J1**	**J2**	**J3**	**J4**	**J5**	**J6**	**Hand**
nr.1	21.86	0	7.48	3.17	5.44	69.44	4.23	4.11	8.86	4.21	15.81	17.28	64.82	0
nr.2	10.07	0	3.07	3.38	2.76	23.18	2.47	1.41	3.64	1.38	2.10	1.65	10.89	0
nr.3	70.08	0	26.25	18.41	60.93	433.33	33.55	38.96	47.88	3.01	18.74	25.02	174.91	23.68
nr.4	24.11	0	10.89	4.17	19.75	107.11	16.42	14.14	21.51	6.09	16.59	23.07	88.35	1.31
nr.5	17.85	0	6.03	3.07	22.16	57.04	13.07	3.70	6.32	0.46	2.30	2.18	15.20	0
nr.6	18.52	0	7.37	4.04	10.88	41.03	5.67	4.60	5.78	1.16	2.37	4.27	45.73	0
nr.7	16.29	0	12.78	0	9.14	49.26	4.40	4.47	7.05	1.35	9.03	13.36	49.17	1.11
nr.8	27.41	0	10.54	7.31	47.44	80.44	11.94	3.17	4.86	1.36	10.01	10.22	68.31	1.76
nr.9	29.25	0	12.16	8.16	12.75	153.43	8.76	16.61	28.51	7.67	39.47	47.69	112.09	0.02
nr.10	14.11	0	4.03	2.96	6.87	67.36	9.08	3.21	4.93	2.07	6.31	10.57	44.03	0.78
nr.11	10.36	0	2.79	1.71	7.27	40.57	5.15	3.5	4.19	0.70	2.82	5.28	28.54	0.01
nr.12	39.16	0	6.71	7.47	35.49	248.73	33.66	25.26	30.94	6.76	30.49	35.78	137.63	5.26
nr.13	14.26	0	2.86	2.87	6.39	64.82	3.52	4.53	12.10	6.13	9.44	11.95	59.93	1.53
nr.14	25.47	0	10.30	11.72	34.35	119.12	9.71	7.50	15.20	5.22	10.2	9.74	89.67	0.05
nr.15	30.68	0	10.94	7.79	14.22	93.00	16.63	12.01	14.72	1.98	8.64	10.25	77.72	2.15
nr.16	23.94	0	6.15	4.36	19.10	125.42	12.28	5.45	10.47	5.67	11.47	15.47	91.14	0.41
nr.17	9.99	0	3.34	0.82	5.56	30.95	1.17	0.49	3.79	1.56	5.71	7.17	24.62	0
nr.18	34.53	0	14.36	8.56	25.70	228.72	19.79	14.24	24.97	3.05	14.85	16.61	140.37	1.43
nr.19	18.33	0	5.73	3.30	10.94	82.12	5.98	5.17	10.75	3.85	7.69	8.1	72.66	2.21
nr.20	27.81	0	12.15	12.09	16.92	130.90	7.51	10.88	34.94	14.98	41.29	22.25	34.12	0.21

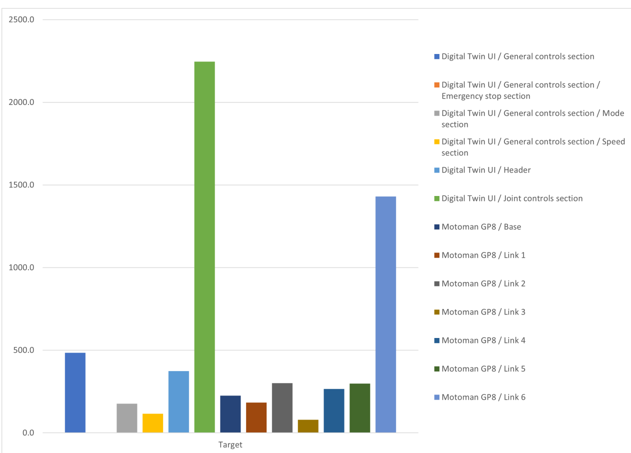

Figure 8. Total seconds spent viewing targets in VR for Group B.

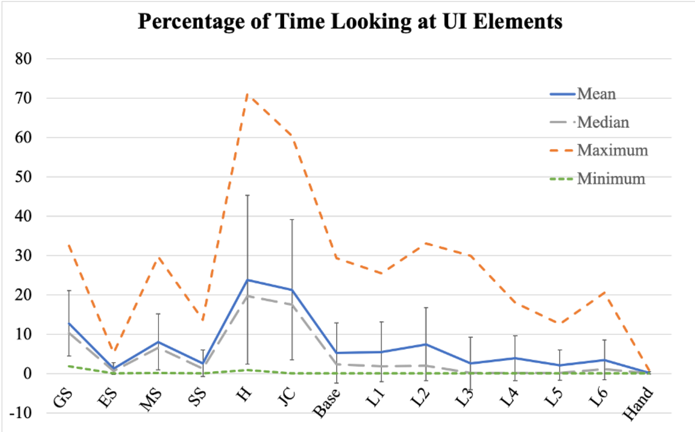

Figure 9. Average operator use of the interface, plotting the mean, median, maximum, and minimum amounts of time spent focusing on different parts of the UI for Group A. The error bars represent a single standard deviation from the mean. The labels on the horizontal axis are those introduced in Table 8.

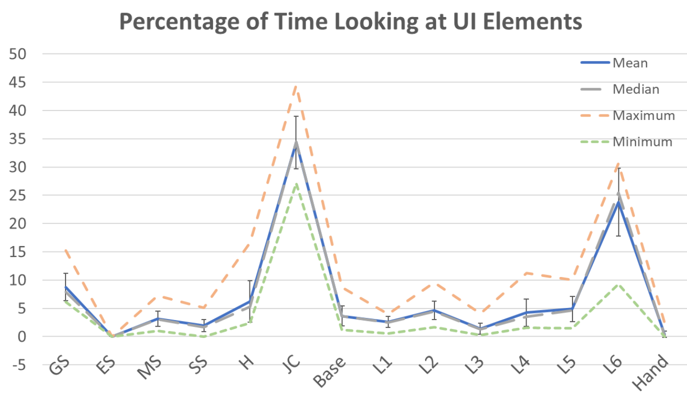

Figure 10. Average operator use of the interface, plotting the mean, median, maximum, and minimum amounts of time spent focusing on different parts of the UI for Group B. The error bars represent a single standard deviation from the mean. The labels on the horizontal axis are those introduced in Table 8.

4.4. Physiological Stress Monitoring

Heart rate data was collected during both physical and virtual trials to assess physiological stress during the test. However, while the results for Group A (Table 10) shows a slight elevation of heart rate during physical trials, this difference is within a single standard deviation as is therefore not significant. For Group B (Table 11) results appear to be even less divergent with nearly identical average heart rate for the physical and virtual robot programming sessions. Due to the relatively low sample size, even with the slightly higher reported range for Group A, the discrepancy is likely the result of a single outlier participant, as the average is within a single standard deviation for both groups. Moreover, the source of the slower heart rate during the virtual trials could not be isolated, as it was not clear if it stemmed from operator comfort during the test, or merely the order of experimentation. This highlights yet another factor that needs addressing in future experiments. Likewise, contrasting these results with the subjective reporting in Section 4.2, there is no clear correlation between heart rate and the perception of effort.

Table 10. Heart rate values during physical and virtual robot programming for Group A.

Process	Maximum (BPM)	Minimum (BPM)	Average (BPM)	SD
Physical robot programming	117	75	90.05	11.99
Virtual robot programming	105	76	85.05	7.53

Table 11. Heart rate values during physical and virtual robot programming for Group B.

Process	Maximum (BPM)	Minimum (BPM)	Average (BPM)	SD
Physical robot programming	99	83	90.55	4.27
Virtual robot programming	100	79	90.1	6.66

5. Discussion

5.1. Advantages and Limitations of the DT System

Results highlighted in Section 4 imply that interacting and controlling a real robot with a traditional tech pendant is largely comparable to VR DT interface control. The VR interface shows better performance overall in terms of time spent in placing the objects with a relevant lower average time after user acquaintance with the virtual environment. Nevertheless, the VR scenario creates more anxiety, and is more demanding on the operator both mentally and physically while not considerably effecting the physiological stress level. The DT system shows promising results in terms of acceptability by the user and overall task execution performance supporting the belief that VR can be a valuable alternative to traditional robot programming interfaces.

Eye-tracking results show that user attention is more frequently directed to the main robot VR UI while not so often to the robot twin. This could probably be due to the perceived safety of the environment. With no real robot moving and being a possible source of danger, the operator might have been able to focus on the interface more without checking the robot position. This hypothesis needs to be confirmed as the precision of the attention system could be also a cause for the collected data set. Attention tracking results could also be influenced by the type of interface interaction in VR. The virtual UI needs to be constantly looked at to be able to use VR pointer selection and interaction as shown in Figure 2. A comparison between eye-tracking attention values in physical and DT trials could clarify the causes of this type of behavior in the VR DT scenario. Furthermore, the current DT control panel does not provide an option to use Inverse Kinematics (IK) when setting the robot positions. Using IK in real-time could speed up the process of working with the digitized robot and bring it on the same feature level as the real machine. Another limitation of the study is the evaluation of familiarity with HMDs, navigation, and interaction in virtual reality. Considering the positive results of task performance in users that took the test in VR twice, it would be informative to understand if improvements were produced by familiarity with the UI, or with the VR interaction and navigation system in general.

5.2. Potential Future Developments (Based on the Findings)

Throughout the analysis of the results, many new questions arose as anomalies and inconsistencies manifested. Moreover, the stated hypothesis in Section 1 could not be fully accepted and rejected and more additional studies should be performed for confirmation of it. Potential future developments of this work can possibly include running the experiment with DT in coupled mode. This could help to determine if the virtual UI allows for better performance than the teach pendant. This could support the design, implementation, and evaluation of different virtual user interfaces for the same robot but customized to different use cases and manufacturing tasks. Furthermore, an advanced programming VR UI for expert users could facilitate the comparison between the two interfaces among proficient users. As mentioned previously, the user’s level of acquaintance with VR interfaces could be considered in advance. Results show that performance time values were considerably lower after a first try of the system. Including eye-tracking in the real robot control scenario and compare results with the collected data from the VR interface would allow for further attention analysis between the teach pendant and virtual UI. Integrating and improving the proposed assessment methodology with virtual reality tools and hardware would allow for the implementation of a fast assessment tool for DT VR interfaces.

Given the results of the eye tracker implementation, it is clear the tracking solution does provide useful information in terms of accuracy, but has insufficient precision. Future efforts will attempt to eliminate the limitations of the current eye tracking approach. The proposed approach as described in Section 4.3 is planned for future implementations. Similarly, as discussed in Section 4.4, heart-rate as a surrogate for stress is currently inconclusive when contrasted with the NASA-TLX survey results. Future efforts can attempt to factor out possible sources of bias including variations in order of operations, proximity to the robot, and tasks.

6. Conclusions

Results gathered during the experiments are pretty promising in blurring the line between the virtual and physical experience of human operators when interacting with industrial robots. The collected data shows no relevant difference in operator journey between the two experimental setups. Moreover, there was no significant difference between group A and group B, which can state that the counterbalance reached its purpose in making the experimental flow more general, and the order of experimental flow did not affect the main flow-only the time of performance with group B on the physical robot was slightly different. The proposed system should be developed further, made more interactive, adapted and integrated to more use-case scenarios. Future work will try to improve the eye-tracking system setup and evaluation for a more efficient assessment of the focus of attention. Nonetheless, it can be stated that the aim of this paper was fulfilled, and research is ready to be continued in preparation for the verification and validation of standardized test methods for DT in HRI.

7. Disclaimer

Certain commercial equipment, instruments, or materials are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by Tallinn University of Technology or the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Author Contributions: Conceptualization, V.K.; Data curation, S.L.P.; Formal analysis, J.A.M. and S.L.P.; Funding acquisition, T.O.; Investigation, M.A. and M.S.; Methodology, J.A.M. and M.S.; Project administration, V.K. and T.O.; Resources, V.K. and Y.B.; Software, Y.B.; Validation, J.A.M., M.A. and S.L.P.; Visualization, M.A. and Y.B.; Writing—original draft, V.K. and S.L.P.; Writing—review & editing, J.A.M. and T.O. All authors have read and agreed to the published version of the manuscript.

Funding: The research was conducted using the Smart Industry Centre (SmartIC) core facility funded by the Estonian Research Council grant TT2. Additionally, supported by the EU Regional Development Fund and by the Estonian Research Council under grant PSG453 "Digital twin for propulsion drive of an autonomous electric vehicle".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

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

Acknowledgments: The research was conducted using the Smart Industry Centre (SmartIC) core facility funded by the Estonian Research Council grant TT2. Additionally, supported by the EU Regional Development Fund and by the Estonian Research Council under grant PSG453 "Digital twin for propulsion drive of an autonomous electric vehicle". Additionally, authors are grateful to student Marietta Gavriljuk, who did aid with the experimental process.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, M.; Fang, S.; Dong, H.; Xu, C. Review of digital twin about concepts, technologies, and industrial applications. *J. Manuf. Syst.* **2020**, *58* , 346–361. [CrossRef]
2. Villani, V.; Pini, F.; Leali, F.; Secchi, C. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics* **2018**, *55*, 248–266. [CrossRef]
3. Berg, J.; Lu, S. Review of Interfaces for Industrial Human-Robot Interaction. *Curr. Robot. Rep.* **2020**, *1*, 27–34. [CrossRef]
4. Krupke, D.; Steinicke, F.; Lubos, P.; Jonetzko, Y.; Görner, M.; Zhang, J. Comparison of multimodal heading and pointing gestures for co-located mixed reality human–robot interaction. In Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2018; pp. 1–9.
5. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Adv. Psychol.* **1988**, *52*, 139–183.
6. Brooke, J. SUS-A "quick and dirty" usability scale. *Usability Eval. Ind.* **1996**, *189*, 4–7.
7. Hassenzahl, M.; Burmester, M.; Koller, F. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Mensch & Computer 2003*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 187–196.
8. Whitney, D.; Rosen, E.; Phillips, E.; Konidaris, G.; Tellex, S. Comparing robot grasping teleoperation across desktop and virtual reality with ROS reality. In *Robotics Research*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 335–350.
9. Tadeja, S.K.; Lu, Y.; Seshadri, P.; Kristensson, P.O. Digital Twin Assessments in Virtual Reality: An Explorational Study with Aeroengines. In Proceedings of the 2020 IEEE Aerospace Conference, Missoula, MT, USA, 7–14 March 2020; pp. 1–13.
10. Kennedy, R.S.; Lane, N.E.; Berbaum, K.S.; Lilienthal, M.G. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.* **1993**, *3*, 203–220. [CrossRef]
11. Rheinberg, F.; Engeser, S.; Vollmeyer, R. Measuring components of flow: The Flow-Short-Scale. In Proceedings of the 1st International Positive Psychology Summit, Washington, DC, USA, 3–6 October 2002.
12. Schubert, T.; Friedmann, F.; Regenbrecht, H. Igroup presence questionnaire. *Teleoperators Virtual Environ.* **2001**, *41*, 115–124.
13. Laaki, H.; Miche, Y.; Tammi, K. Prototyping a Digital Twin for Real Time Remote Control Over Mobile Networks: Application of Remote Surgery. *IEEE Access* **2019**, *7*, 20325–20336. [CrossRef]
14. Concannon, D.; Flynn, R.; Murray, N. A Quality of Experience Evaluation System and Research Challenges for Networked Virtual Reality-Based Teleoperation Applications. In Proceedings of the 11th ACM Workshop on Immersive Mixed and Virtual Environment Systems; Association for Computing Machinery, MMVE'19, New York, NY, USA, 18 June 2019; pp. 10–12. [CrossRef]
15. Matsas, E.; Vosniakos, G.C.; Batras, D. Effectiveness and acceptability of a virtual environment for assessing human–robot collaboration in manufacturing. *Int. J. Adv. Manuf. Technol.* **2017**, *92*, 3903–3917. [CrossRef]
16. Oyekan, J.O.; Hutabarat, W.; Tiwari, A.; Grech, R.; Aung, M.H.; Mariani, M.P.; López-Dávalos, L.; Ricaud, T.; Singh, S.; Dupuis, C. The effectiveness of virtual environments in developing collaborative strategies between industrial robots and humans. *Robot. Comput.-Integr. Manuf.* **2019**, *55*, 41–54. [CrossRef]
17. Sievers, T.S.; Schmitt, B.; Rückert, P.; Petersen, M.; Tracht, K. Concept of a Mixed-Reality Learning Environment for Collaborative Robotics. *Procedia Manuf.* **2020**, *45*, 19–24. [CrossRef]
18. Yap, H.J.; Taha, Z.; Dawal, S.Z.M.; Chang, S.W. Virtual reality based support system for layout planning and programming of an industrial robotic work cell. *PLoS ONE* **2014**, *9*, e109692.
19. Malik, A.A.; Masood, T.; Bilberg, A. Virtual reality in manufacturing: Immersive and collaborative artificial-reality in design of human–robot workspace. *Int. J. Comput. Integr. Manuf.* **2020**, *33*, 22–37. [CrossRef]
20. Pérez, L.; Rodríguez-Jiménez, S.; Rodríguez, N.; Usamentiaga, R.; García, D.F. Digital Twin and Virtual Reality Based Methodology for Multi-Robot Manufacturing Cell Commissioning. *Appl. Sci.* **2020**, *10*, 3633. [CrossRef]
21. Peruzzini, M.; Grandi, F.; Pellicciari, M. Exploring the potential of Operator 4.0 interface and monitoring. *Comput. Ind. Eng.* **2020**, *139*, 105600. [CrossRef]
22. Karwowski, W. The OCRA Method: Assessment of Exposure to Occupational Repetitive Actions of the Upper Limbs. In *International Encyclopedia of Ergonomics and Human Factors-3 Volume Set*; CRC Press: Boca Raton, FL, USA, 2006; pp. 3337–3345.
23. Havard, V.; Jeanne, B.; Lacomblez, M.; Baudry, D. Digital twin and virtual reality: A co-simulation environment for design and assessment of industrial workstations. *Prod. Manuf. Res.* **2019**, *7*, 472–489. [CrossRef]
24. McAtamney, L.; Corlett, N. Rapid upper limb assessment (RULA). In *Handbook of Human Factors and Ergonomics Methods*; CRC Press: Boca Raton, FL, USA, 2004; pp. 86–96.
25. Rosen, E.; Whitney, D.; Phillips, E.; Chien, G.; Tompkin, J.; Konidaris, G.; Tellex, S. Communicating robot arm motion intent through mixed reality head-mounted displays. In *Robotics Research*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 301–316.
26. Moniri, M.M.; Valcarcel, F.A.E.; Merkel, D.; Sonntag, D. Human gaze and focus-of-attention in dual reality human–robot collaboration. In Proceedings of the 2016 12th International Conference on Intelligent Environments (IE), London, UK, 14–16 September 2016; pp. 238–241.

27. Wassermann, J.; Vick, A.; Krüger, J. Intuitive robot programming through environment perception, augmented reality simulation and automated program verification. *Procedia CIRP* **2018**, *76*, 161–166. [CrossRef]
28. Vosniakos, G.C.; Ouillon, L.; Matsas, E. Exploration of two safety strategies in human–robot collaborative manufacturing using Virtual Reality. *Procedia Manuf.* **2019**, *38*, 524–531. [CrossRef]
29. Maragkos, C.; Vosniakos, G.C.; Matsas, E. Virtual reality assisted robot programming for human collaboration. *Procedia Manuf.* **2019**, *38*, 1697–1704. [CrossRef]
30. Manou, E.; Vosniakos, G.C.; Matsas, E. Off-line programming of an industrial robot in a virtual reality environment. *Int. J. Interact. Des. Manuf.* **2019**, *13*, 507–519. [CrossRef]
31. Pérez, L.; Diez, E.; Usamentiaga, R.; García, D.F. Industrial robot control and operator training using virtual reality interfaces. *Comput. Ind.* **2019**, *109*, 114–120. [CrossRef]
32. Ostanin, M.; Klimchik, A. Interactive Robot Programming Using Mixed Reality. *IFAC-PapersOnLine* **2018**, *51*, 50–55. [CrossRef]
33. Nathanael, D.; Mosialos, S.; Vosniakos, G.C. Development and evaluation of a virtual training environment for on-line robot programming. *Int. J. Ind. Ergon.* **2016**, *53*, 274–283. [CrossRef]
34. Ong, S.; Yew, A.; Thanigaivel, N.; Nee, A. Augmented reality-assisted robot programming system for industrial applications. *Robot. Comput.-Integr. Manuf.* **2020**, *61*, 101820. [CrossRef]
35. Burghardt, A.; Szybicki, D.; Gierlak, P.; Kurc, K.; Pietruś, P.; Cygan, R. Programming of Industrial Robots Using Virtual Reality and Digital Twins. *Appl. Sci.* **2020**, *10*, 486. [CrossRef]
36. Kuts, V.; Otto, T.; Tähemaa, T.; Bondarenko, Y. Digital twin based synchronised control and simulation of the industrial robotic cell using virtual reality. *J. Mach. Eng.* **2019**, *19*, 128–144. [CrossRef]
37. Marvel, J.A.; Bagchi, S.; Zimmerman, M.; Antonishek, B. Towards Effective Interface Designs for Collaborative HRI in Manufacturing: Metrics and Measures. *ACM Trans. Hum.-Robot. Interact.* **2020**, *9*, 1–55. [CrossRef]
38. Bartneck, C.; Kulic, D.; Croft, E.A. Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots. In Proceedings of the Metrics for Human-Robot Interaction, ACM/IEEE International Conference on Human-Robot Interaction, Amsterdam, The Netherlands, 12–15 March 2008; ACM: Amsterdam, The Netherlands, 2008.
39. *ISO 25010*; Systems and Software Engineering-Systems and Software Quality Requirements and Evaluation (SQuaRE)-System and Software Quality Models. International Organization for Standardization: Geneva, Switzerland, 2011.
40. Kuts, V.; Modoni, G.E.; Otto, T.; Sacco, M.; Tähemaa, T.; Bondarenko, Y.; Wang, R. Synchronizing physical factory and its digital twin throughan iiot middleware: A case study. *Proc. Est. Acad. Sci.* **2019**, *68*, 364–370. [CrossRef]
41. Kuts, V.; Cherezova, N.; Sarkans, M.; Otto, T. Digital Twin: Industrial robot kinematic model integration to the virtual reality environment. *J. Mach. Eng.* **2020**, *20*, 53–64. [CrossRef]

IV publication

IV

S.L. Pizzagalli, Y. Bondarenko, V. Kuts, E. O'Connell, N. Murray, and T. Otto. Ros-based augmented and virtual reality path planning interface for industrial robotic arms: a preliminary assessment. In *ACM, the 6th International Conference on Robot Systems and Applications (ICRSA), 2023*. ACM, 2024

ROS-based Augmented and Virtual Reality path planning interface for industrial robotic arms: a preliminary assessment

Simone Luca Pizzagalli
Tallinn University of Technology, Dept. of Mechanical and Industrial Engineering
Tallinn, Estonia
simone.pizzagalli@taltech.ee

Yevhen Bondarenko
Tallinn University of Technology, Dept. of Mechanical and Industrial Engineering
Tallinn, Estonia
yevhen.bondarenko@taltech.ee

Vladimir Kuts
Tallinn University of Technology, Dept. of Mechanical and Industrial Engineering
Tallinn, Estonia
vladimir.kuts@taltech.ee

Eoin O'Connell
University of Limerick, Dept. Of Electronics and Computer Engineering
Limerick, Ireland
eoin.oconnell@ul.ie

Niall Murray
Technological University of the Shannon, Dept. of Electronics and Informatics
Athlone, Ireland
nmurray@research.ait.ie

Tauno Otto
Tallinn University of Technology, Dept. of Mechanical and Industrial Engineering
Tallinn, Estonia
tauno.otto@taltech.ee

ABSTRACT

Optimised human machine interfaces for multi-robot systems are essential for human in the loop in cyber-physical production lines and collaborative systems. The speed of changeover and customized production processes together with the need for straightforward easy and user-friendly human-machine interaction methods demands natural and adaptable interfaces. They should be based on flexible software cores and packages that ease and speed up the development processes. This also applies to laboratory testing and assessment in the academic field, in particular when it comes to the deployment of virtual or augmented reality (AR/VR) interfaces. The standardization of human-robot interaction including control methods in the extended reality domain is a work in progress. It needs broader assessment and clearer metrics to realise efficient and reliable tools. This work presents an Extended Reality (XR) user interface for the control and teleoperation of industrial robots. The systems allows the fast integration of the digital twins of robotic arms and path planning interface in AR and VR using Robot Operating System and Unity. Furthermore, a design-of-experiment involving two different robots (ABB IRB 1200 and ABB IRB 1600) in the two geographically distributed locations is proposed along with some preliminary experimental results.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; **Usability testing**; **Mixed / augmented reality**; **Virtual reality**; • **Computer systems organization** → **Robotic control**.

KEYWORDS

Human Robot Interaction, Design and assessment methods, Augmented Reality, Virtual Reality, User studies, Human Factors

ACM Reference Format:
Simone Luca Pizzagalli, Yevhen Bondarenko, Vladimir Kuts, Eoin O'Connell, Niall Murray, and Tauno Otto. 2023. ROS-based Augmented and Virtual Reality path planning interface for industrial robotic arms: a preliminary assessment. In *2023 the 6th International Conference on Robot Systems and Applications (ICRSA) (ICRSA 2023), September 22–24, 2023, Wuhan, China.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3655532.3655538

1 INTRODUCTION

Recent years have seen the evolution and pervasive adoption of cyber-physical systems (CPS) [16] in many aspects of everyday life. The technological backbone of these systems which integrate machine computational power with real-world entities through data collection, exchange, and processing, has expanded tremendously. Widespread high-speed connectivity, digital simulation tools, extended reality (XR), the metaverse, and advanced interfaces, to name a few, boosted the adoption of specific technologies in the industrial domain, towards higher efficiency, flexibility, customization, and sustainability. The integration of digital twins (DT) of production systems [4], artificial intelligence (AI), dynamic immersive interaction interfaces, and pervasive data collection, pushed the industrial paradigm to move from pure quantity-driven automation, what is known as Industry 4.0, to collaboration between operators and machines. This change emerged and developed in a novel industrial paradigm that is in constant evolution. Industry 5.0 (I5.0) constitutes the current convergence point of human skills and collaborative robotic support within a scenario driven by customization and integration of the digital domain, physical world, and human factors. This scenario has several names and definitions [14] and it is certainly expanding in fields outside just manufacturing.

A human centred approach is essential [25] to realise technologies that empower operators. Such technologies can support decision-making, and adapt to needs and requirements, constituting what

ICRSA 2023, September 22–24, 2023, Wuhan, China

ACM ISBN 979-8-4007-0803-9/23/09
https://doi.org/10.1145/3655532.3655538

is called Human Cyber-Physical Systems. DT and XR are tools and techniques which put the operator at the center of this digital/physical merge providing interaction visualization and awareness also in teleoperation scenarios [17]. Nevertheless, there is a need to define borders between the digital and physical aspects and roles of the operator in this new scenario as much as assess the impact of the technologies that allow for symbiotic collaboration on the shop floor. Moreover, it is necessary to establish some practices of integration of new machinery and the development of custom interfaces in the DT XR domain. These should be based on frameworks and methodologies that would be flexible, adaptable, easy to use, modular, and human-centric [7]. In this work, we present a DT XR User Interface (UI) for generic industrial robot control and teleoperation based on Robot Operative System (ROS) and Unity. The proposed framework allow interoperability and flexibility based on a well-established standard while easing the design of immersive or augmented interfaces for Human-Robot Interaction even with older machinery. The study provides a design of experiment for the assessment of the interfaces and interaction methods on a specific task, and their impact on the user. Results of the first experimental sessions are reported together with some preliminary conclusions on system usability, user experience (UX) and task load.

2 RELATED WORKS

A variety of recent works and publications explore the advantages of ROS-based XR UIs in the control and teleoperation of industrial robots. The architecture has been known for some time and employed in remote manipulation on a varied range of tasks, as exemplified in the work by Whitney et al. [26] and including trajectory programming in augmented or virtual reality (AR/VR). Studies like the ones from Shamaine et al. [22] or Piccinelli et al. [21] make use of Unity-based XR applications which leverage the ROS backend to augment the operator's capabilities and overall programming task efficiency. Flexibility requirements and efficient task allocation, or changeovers of industrial systems constitute one of the many use cases benefiting from the features of XR UIs. Togias et al. [24] present an immersive environment for remote robot control and repurposing of an assembly line or production station. While providing a valuable use case description of the ROS-based system, the authors don't examine in depth the types of interactions and operator's engagement in the tasks.

The advantages of such applications are nevertheless clear: controlling industrial systems from afar, being able to include different agents in the same control interface, direct feedback to the operator, or advanced techniques such as collision avoidance, are all characteristics that add up to the concept of user-centered CPSs. In this respect, Yun at al. [27] present an extensive XR-based UI framework aimed at the configuration and design of systems integrating human work and automation, proposing several use cases and tools supporting the augmentation of operators' skills and abilities. Based on these examples, fast and reliable connectivity is fundamental, an example of 5G-supported CPS is demonstrated in [10], as much as the way the operator interacts with the virtual agents, being digital shadows or twins, and how the XR interface is capable of involving the human in a realistic, yet efficient interaction experience. This aspect is paramount and tacked in a few recent studies. Direct hand manipulation seems to be the trending method with examples using both AI-based custom tracking and pose estimation solutions [19] and methods that are delivered with the latest AR/VR headsets, such as in Ostanin et al. [20] and many of the previously mentioned works for path planning tasks. Hand manipulation is often associated with advanced environmental reconstruction and haptic feedback, which is still far to be fully integrated into these types of interfaces. An example is given in the study by Lee et al. [13].

Considering the variety of use cases and technologies that aim at integrating human intelligence and machine efficiency in the modern industry there is still a lack of well-defined and reliable metrics and methods. A few studies propose an evaluation of different robot control methodologies such as in Kuts et al. [11], some of them utilizing similar software architectures and XR tools, such as Hetrick et al. [9] or Solanes et al. [23]. Results of the comparative study provided by Naceri et al. [18] show, for instance, how VR UI solutions for the specific use case, are still slow and demanding on the operator.

The presented system provides a simple, standard solution to easily set up XR UI for robot programming with different machines. The goals are scalability and simple remote accessibility and setup through docker containers with the possibility of remotely deploying the system in real industrial scenarios. Paramount importance is given to the interaction methods within the XR UIs which should overcome difficult physical interface setups and reduce the overall system learning rate. The main objective of this study, which is connected to the industrial paradigm of I5.0 and not fully acknowledged in existing literature, is to address human factors and assess the efficiency of XR UIs in simple robot control and teleoperation tasks. As much as cyber-physical systems are moving back towards human operators, these should also be included in the design process while assessing what kind of impact these systems have on them and how efficient and effective new interfaces are compared to the traditional ones.

3 SYSTEM ARCHITECTURE

3.1 Hardware

The hardware equipment used for the XR visualization consists of Meta Quest 2 and HoloLens 2 headsets. Both can operate as standalone devices without the need for any cabled connection to a workstation. Each headset has environment and hand-tracking capabilities, based on infrared cameras, which are leveraged for UI interactions as explained in the following sections. The use case we present in this paper, and the consequent design of experiment, are based on two six-axis articulated robots from ABB, namely ABB IRB 1600 and ABB IRB 1200. As already mentioned, the system is meant to be extensible to other machines and robots while using the same software and interface infrastructure.

3.2 Software

As mentioned before, the system is based on two main software components: ROS, used to control the robot, and Unity, used for visualization and interactions. The foremost goal for the software architecture was to make it easy to integrate with different models of industrial robots, while requiring minimal changes. For this reason, in the ROS part of the system we opted for ROS-Industrial (ROS-I) [1] standard combined with MoveIt motion planning framework [6]. ROS-I defines a unified way of providing industrial robot packages for ROS, which include visual and collision geometry of the robots, their ROS control files and MoveIt configurations. Relying on this standard allows our system to work with any ROS-I-supported robot model while preserving the same programming interface and robot control functionality available to the connected Unity application.

Control of the physical robot is handled fully by ROS. The Unity application communicates with ROS nodes over the network to visualize the up-to-date robot poses to the operator and send motion commands to the real hardware. The Unity side of the software architecture is shown in Fig. 1

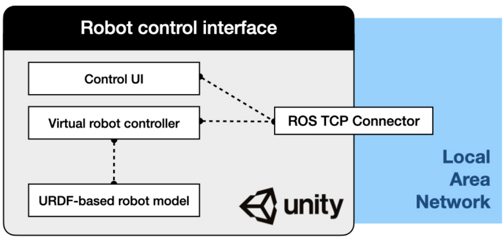

Figure 1: Robot path planning and control interface architecture, Unity.

Communication with ROS is implemented based on the Unity's Robotics Hub project [2], with extra scripts implemented to handle MoveIt-specific topics and message types. Using standard MoveIt control topics in our Unity scripts creates an abstraction layer for control logic, which makes the Unity application independent from the control scripts of the specific robot connected on the ROS side. The robot control server from the ROS side is shown in Fig. 2

3D model of the robot used for visualization in Unity is imported from a Universal Robot Description Format file, which comes from the corresponding ROS robot package compatible with ROS-I standard. Such implementation makes the import process of new robot models trivial and allows to keep using the same AR and VR interfaces even if another robot model gets connected to the system, thanks to the abstraction layer provided by our scripts.

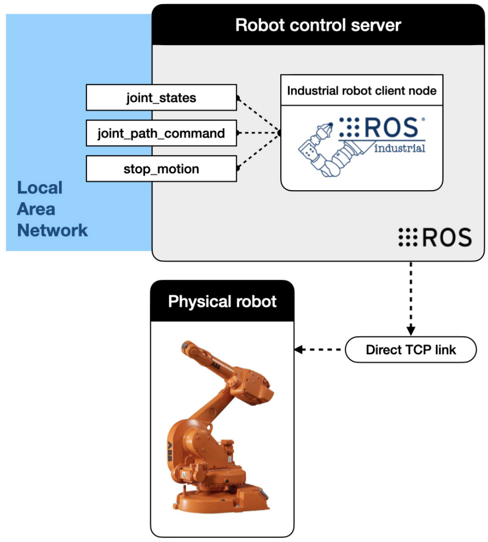

Figure 2: Robot control server architecture, ROS.

Figure 3: ABB DSQC679 teach pendant

4 USER INTERFACES

4.1 Native teach pendant

The native teach pendant for the ABB IRB 1600 and ABB IRB 1200 robotic arms is shown in Fig. 3, model DSQC679. This is the same for both ABB robots used in our work. The robotic arms can be manipulated directly through teach pendant interaction as provided by the producer.

The procedure required to save a path goes through several preliminary navigation steps necessary to access the correct routines location and set up the appropriate parameters for the robot joint speed and movements. These actions happen through the teach pendant monitor navigation, capable of touch or pen input, physical interface buttons and the six axes joystick.

4.2 Interaction methods

One of the goals of the application is to ease user interactions within the virtual and augmented reality interface and support the user experience while performing the task, so that operator can concentrate on the task goals without spending time dealing with

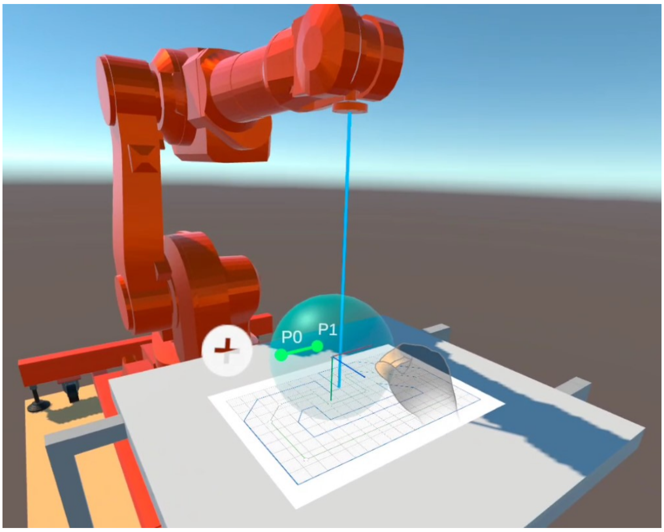

Figure 4: ABB1600 VR setup with hand manipulation of the path planning tool. The first two path points are visible in green in the figure.

the interface itself. Previous experiments run on similar XR interfaces, compared several interaction types for immersive virtual UI, with the purpose of achieving a more natural workflow while controlling industrial robots. These studies focused on immersive VR interfaces while adopting different controller-based UI interactions including distance laser pointer interaction, joystick navigation, and direct interaction with UI elements, mediated, also this time, by the headsets device controllers. In this study, we decided to work in the same direction but opt for full hand manipulation and interaction of the UI components. For this purpose, we adopted Meta Quest 2 and HoloLens support for hand tracking. Both devices allow tracking hand and finger position and intuitive interactions with holograms, virtual objects, and interfaces. The VR setup with hand avatar manipulation of the path planning tool is shown in Fig.4.

HoloLens 2 and Meta Quest 2 leverage on inside-out camera tracking method by using embedded infrared cameras to determine the position and orientation of the hands, and objects in the surrounding environment. This system has become more and more popular as it needs less hardware and allows the user wearing the head-mounted display (HMD) to move freely in space. The UI for both immersive and augmented interfaces is structured in the same way, while using two different software development kits. For HoloLens 2 we employed the Microsoft Mixed Reality Toolkit (MRTK) [15] which provides an extensive collection of tools to create and customize augmented UIs from scratch. In our system, we included basic components such as buttons and checkboxes. The whole UI can be also moved and rotated, as any other 3D object, thanks to the bounds control component provided with MRTK. In the case of Meta Quest 2, we adopted the native Unity UI package and integrated hand interaction through the script components provided with the Oculus Integration Software Development Kit (SDK) for Unity. All XR buttons and UI components can be interacted through an index finger poke gesture as a common physical hardware interface element.

4.3 Interface functions

The interface is structured as follows. The top menu bar allows checking on the ROS server connection and quitting the application through the dedicated button. The Server Connection background turns red or green based on the actual state of the connection. The first section allows to select an existing path, delete it, or create a new one. The path selection menu varies in AR and VR applications. In the former, the existing paths are visualized in a separate tab called Scrolling Collection, which is a component included in the MRTK SDK, and provided as buttons. In the latter, the paths are visualized as options in a drop-down menu. Each created path is saved on a hardware local folder in .json format. The file includes information about the path name, total number of points, pose, position and rotation of each created point in the path. The second menu section contains two buttons dedicated to controlling and actuating the robot. The first allows setting the position of the robot to home, which is normally with the end effector looking downwards towards the working surface. The second button allows the robot to follow the created routine. The last interface section includes a button used to create path points and a checkbox that allows locking the vertical position of the path planner manipulator tool. This should facilitate the user in creating a path consistent with the example maze provided and as precise as possible. An

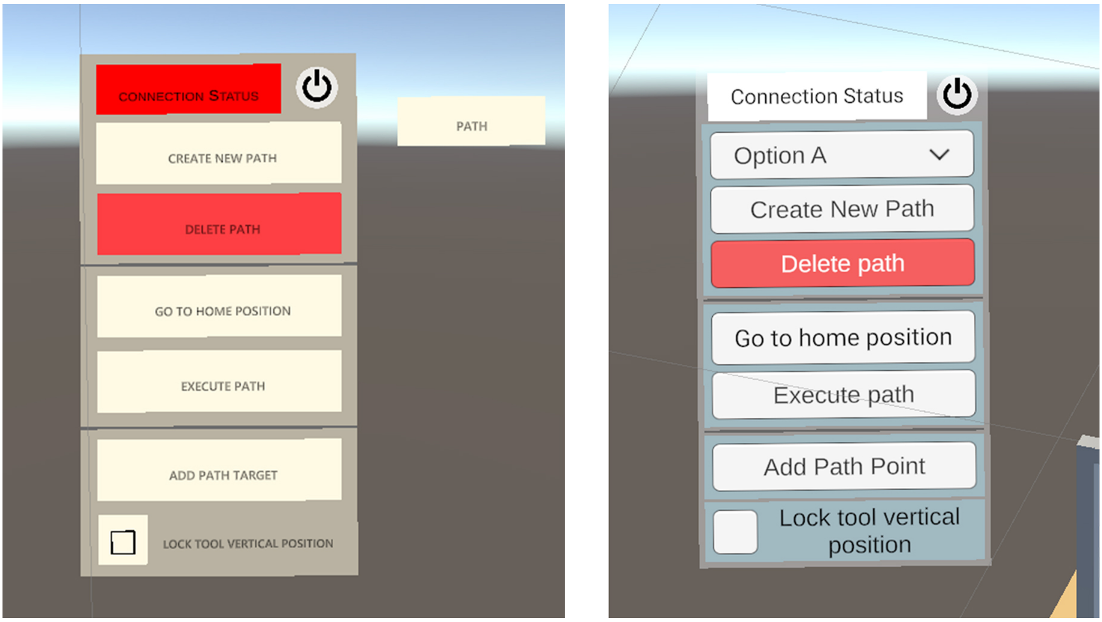

Figure 5: XR UI functions and sections (AR on the left, VR on the right)

overview of the UI controls in both AR and VR scenario are shown in Fig. 5.

4.4 Path planning tool

The path planning tool provided in both augmented and virtual reality scenarios is presented as a sphere that has a few properties and components facilitating the creation of the robot paths. The sphere itself is transparent, to avoid the occlusion of the robot, path, or the example maze. Cartesian axes and origin point references, are included in the tool and allow the correct positioning and orientation during the task. The center point is specifically meant to locate the newly spawned path points in the correct world space position. A UI button is provided adjacent to the sphere, allowing for the rapid instancing of path points once the path has been initialized in the main interface. In both scenarios, the path planning tool can be grabbed through hand gestures provided by both Oculus and MRTK SDKs. Fig. 6 shows the AR setup with the ABB1600 robot. The figure shows the list of available paths, the main UI button and functions, together with the planning tool and separated button to create new path points. This interaction is meant to be natural and intuitive and allow for the rapid creation of complex path routines.

5 DESIGN OF EXPERIMENTS AND SYSTEM ASSESSMENT

Before initiating a full experimentation and interface comparison the system has been tested with both robots in the two research laboratory premises. These tests were only meant to estimate the basic robot teleoperation functions, hand interaction, and UI elements. After the first tests the stability of target tracking for the Augmented Reality interface has been addressed for the optimal user experience of this type of robot control. Laboratory lighting conditions and floor surface reflections were modified to improve camera tracking by positioning the target image on a matte surface improving the targeting quite noticeably. For the sake of assessing the efficiency, effectiveness, user experience, and operator workload, we developed an experimental protocol and design of experiments that include both robots in a similar task. Two sets of experiments are planned to involve 50 experienced and non-experienced users in a simple path-planning task. The experiments aim at comparing interaction with the robot's native teach pendant, the immersive VR interface running on Meta Quest 2, and the AR application running on HoloLens 2. The setup for this last headset is shown in Fig. 7.

During each session the users are asked to follow an exemplifying path, create way-points in the correct positions, and finally run the resulting robot routine. A laser pointer, both real and virtual, guides the users in achieving the task in the most effective way. Experimental sessions are counterbalanced among the different types of interfaces (AR, VR and, Teach Pendant). To be able to statistically analyze within-group test results, we decided to collect both qualitative and quantitative data during the different experimental sessions. These include the time spent in accomplishing the path planning task and biometrics data (heart rate) for the estimation of stress level and workload. We also include three different well-known and validated questionnaires to assess respectively user

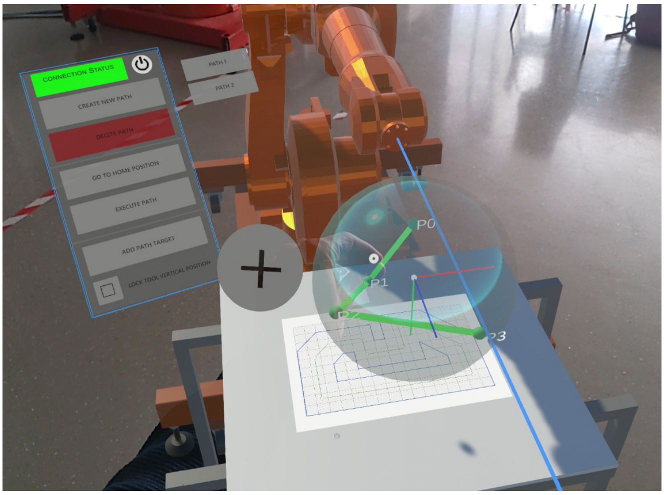

Figure 6: The AR setup with ABB1600 robot with UI and path planning tool. Path points and the current loaded path are visible in green.

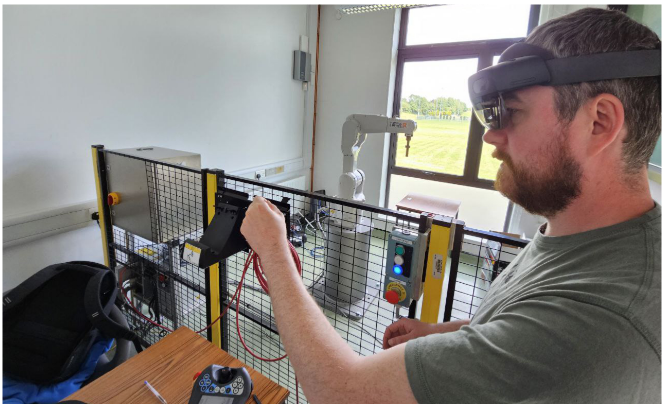

Figure 7: The AR setup with ABB1200 robot. The operator is manipulating the augmented path planning tool.

experience, User Experience Questionnaire (UEQ), task workload, Nasa Task Load Index (TLX), and usability, System Usability Scale (SUS). The experimental protocol, including the data collected in each session, is illustrated in Fig. 8

6 PRELIMINARY RESULTS

6.1 Evaluation metrics

As already mentioned, quantitative and qualitative data are collected during the experimental sessions. In particular time spent in completing the task and heart rate to detect stress levels while interacting with the system. NASA TLX [8] questionnaire is used to assess workload during the task performance for each user with each different UI. The results of this questionnaire are scored on a scale from 0 to 100 on six items: mental demand, physical demand, temporal demand, performance, effort and frustration level. We report here the raw unweighted results for each UI. Perceived task workload should be interpreted based on the values starting from 0, Very Low, to 100, Very High.

The System Usability Scale [5] is adopted to detect the general usability of the system. This questionnaire is based on ten items scored on a five points Likert scale and evaluating the global opinion of the operator on system usability (effectiveness, efficiency and satisfaction). The tool scores the final results on a scale from 0 to

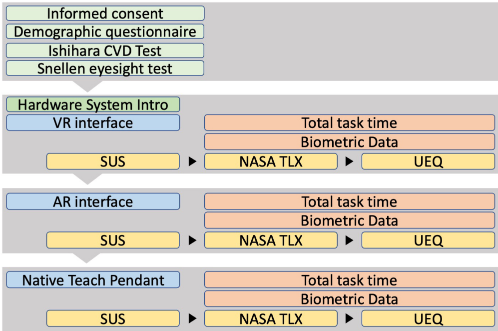

Figure 8: Experimental protocol diagram

100. Figure 9 shows the usability ranking based on the outcome scores.

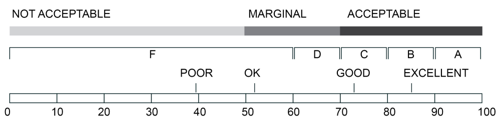

Figure 9: SUS ranking scale

The User Experience Questionnaire [12] is assessing the user experience while interacting respectively with the robot Teach Pendant, the VR and AR interfaces. The short version of the questionnaire has 8 items representing hedonic and pragmatic qualities scored on a 7-point Likert scale and having a positive and negative term. Results are scored on values between -3 and 3 (very bad or very good UX). Values between 0.8 and -0.8 are considerate average or neutral while values above 0.8 represent a good UX rating and values below -0.8 a bad evaluation. Only preliminary results and simple statistical analysis on the raw data are presented in this work as the study is still ongoing. Heart rate data is still not available. Figure 10 shows one of the experiments running on the first group of subjects with the VR setup using Meta Quest2.

6.2 Demographics and preliminary questionnaire

A total of 21 subjects were involved in the first round of experiments taking place at the Technical University of the Shannon Midlands laboratories. The sessions were counterbalanced among the different control and teleoperation methods in groups of 7 subjects per sequence. A total of 13 male and 8 female subjects with an average age of 33 years (SD ±5.8) took part in the experiments. All subjects were normally sighted and with no color blindness.

Almost all of the subjects strongly agree on the potentials and appeal of AR and VR technologies. Average scores on a 5 points Likert scale are respectively 4.6 (SD ±0.5) for the first and 4.5 (SD ±0.6) for the second technology. Expectations on the enjoyability of XR interfaces for teleoperation are also high with an average of 4.5 (SD ±0.5), corresponding to a strong agreement, for both AR and VR. Almost all subjects do not feel scared by operating an industrial robot (average 1.9, SD ±0.9) and disagree with the questionnaire item. Only two persons agreed on the statement while other three were neutral about it. All subjects had previous experience with VR and AR headsets except for one who never used any AR device before. Only 5 subjects had previous experience with industrial robot programming. Table 1 summarizes the basic data on the subjects sample.

Table 1: Characteristics of the sample

Age	33 ± 5.8
Gender	
Male	13
Female	8
Have experience with AR headsets	
Yes	20
No	1
Experience with VR headsets	
Yes	21
No	0
Experience with industrial robot programming	
Yes	5
No	16

6.3 Performance

The Time necessary to perform the task is higher when subjects are operating the robot thought the Teach Pendant with an average of 334.34 seconds and almost half time spent when using the AR (179.57 seconds) and VR (152.76 seconds) UIs. Values for these two interfaces are similar, with higher timing for the AR interface. Standard deviation values for the Tech Pendant are nevertheless high and more experiments are necessary to confirm the results.

Table 2 and Figure 11 shows the results for the task performance (in seconds) for each type of interaction.

6.4 Task load

Based on the preliminary results the VR UI seems to have the least workload on the operator with a score of 26.43 followed by the AR UI with a 32.02. Both values are not very low but more acceptable than the Teach Pendant interface scoring 43.89 overall. The three interfaces show high level of mental demand and general task performance effort (physical and mental work employed to achieve the specific level of performance). The values for the physical and

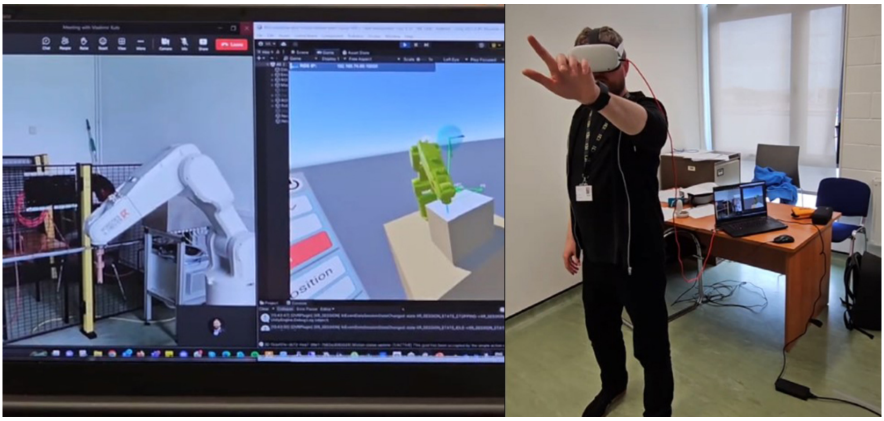

Figure 10: Experiments running in the laboratory. On the screen the real robot is visible together with the VR visualization where the subject is interacting

Table 2: Task time performance (in seconds)

	Mean task time (seconds)	SD
Teach Pendant	334.24	± 162.15
AR UI	179.57	± 58.72
VR UI	152.76	± 55.78

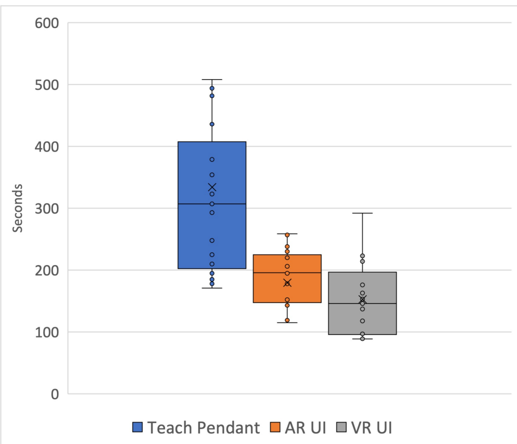

Figure 11: Performance data boxplot

temporal demand in the case of the Tech Pendant interface are also above average with values of 45.95 and 44.29 respectively. Table 3 shows the raw unweighted result of the NASA TLX questionnaire. Figure 12 provides a visual plot of the results for each item and for the overall values.

Table 3: Nasa TLX results

	AR UI	VR UI	Tech Pendant
Mental	38.10	35.00	55.95
Physical	26.43	21.90	45.95
Temporal	29.75	28.10	44.29
Performance	33.00	25.24	33.33
Effort	35.95	28.57	48.81
Frustration	31.90	19.76	35.00
OVERALL	32.02	26.43	43.89

6.5 System Usability and User Experience

The VR UI SUS results show a good usability level with a mean score of 76.54 (SD ± 20.47). This evaluation is confirmed, but lower, for the AR UI with a score of 69.76 (SD ± 21.08). The Tech Pendant scores are not as good showing a poor/ok usability (Mean 50.83 SD ± 24.71). Standard deviation values are nevertheless high indicating a big variability in the population and confirming the necessity of more experiments. Table 4 and Figure 13 shows the result for the SUS questionnaire.

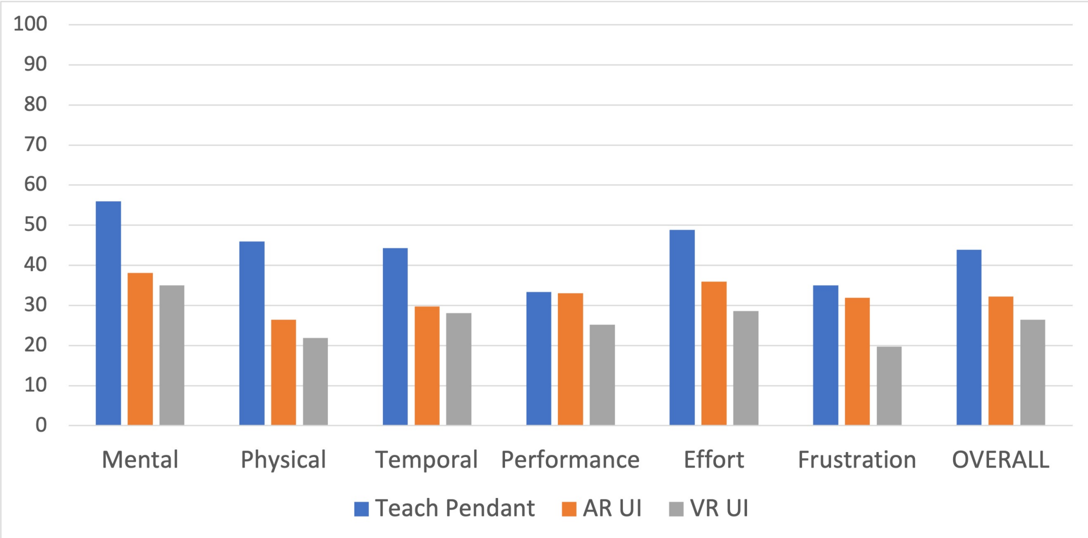

Figure 12: Nasa TLX results

Table 4: SUS questionnaire results

	Mean	SD
Teach Pendant	50.83	± 24.71
AR UI	69.76	± 21.08
VR UI	76.54	± 20.47

The UEQ results show that UX for the VR and AR interfaces is good while results for the Teach Pendant interface are average. Table 5 shows the results for pragmatic quality, hedonic quality and overall UX scores for the three UIs.

Table 5: UEQ questionnaire results

	Pragmatic	Hedonic	OVERALL
Teach Pendant	0.51	0.84	0.68
AR UI	1.55	2.28	1.92
VR UI	1.92	2.21	2.06

Pragmatic quality reports low scores for all the three types of interfaces with the highest one in case of the VR UI (1.92). Hedonic quality scores for the XR interfaces are high, confirming the general positive inclination of the users towards novel system interfaces and interactions. These results are consistent with the good usability

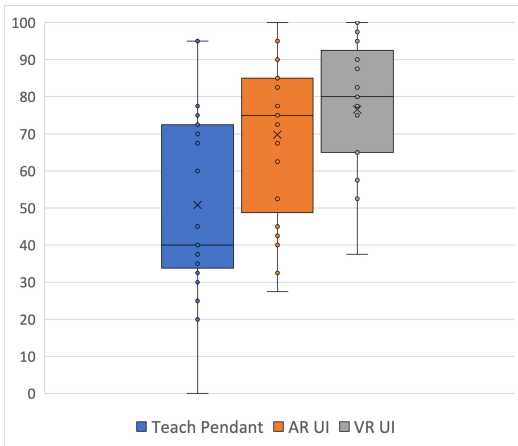

Figure 13: System Usability Scale boxplot

of AR and VR UIs for the specific task. By having a closer look at the questionnaire items average results (not reported in the table above), it is possible to notice that the VR UI scores higher in terms of pragmatic quality, and it is considered easy to use with a score of 2.0 against 1.1 for the AR interface. The lowest items scores for the Teach Pendant show that the interfaces in perceived complicated (mean -0.1) and ordinary (mean -0.2). A summary of UEQ pragmatic

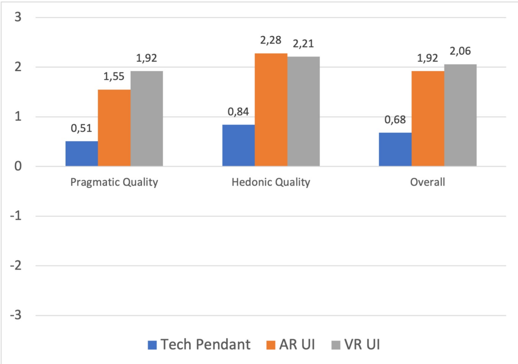

Figure 14: User Experience Questionnaire results

and hedonic quality scores for each interaction method and overall scores are show in Figure 14.

As for other questionnaire results the SD values per item are quite high confirming the necessity of further investigations.

7 CONCLUSIONS AND FUTURE WORKS

The preliminary assessment of the system and simple data analysis on part of the whole experimental sample (21 subjects in total) supports the hypothesis that XR interfaces for robot control and teleoperation are more efficient, easy to use and have minor workload impact on the user than traditional teach pendants. Results of the UEQ and SUS are consistent and showing that the VR UI, followed by the AR UI, are perceived as more efficient and effective in achieving the task and more pleasurable to use. These results are nevertheless limited. More experimental data and proper statistical analysis is required to understand the results related factors and correlations within the population characteristics.

Furthermore the presented system shows high flexibility of using ROS and Unity-based interfaces for the creation of multiple types of interactions aimed at XR human-robot control and teleoperation. The ROS server can be also run inside a Docker [3] container, speeding up the development time, allowing accessibility from anywhere, and keeping the core software always up to date for different integration developments.

The proposed interfaces have several limitations which will be addressed after concluding the experiments. More functions can be included in the UI such as path editing and path point repositioning and the possibility of naming the paths for the specific work. The possibility of rotating the path planning tool will also be considered to allow the creation of more complex and curved paths. Further development of this system will address advanced safety systems for older robots, environmental awareness through depth camera integration, collision avoidance and extended teleoperation testing across the two locations. Furthermore, the hand manipulation interactions might be tested against the already developed interaction VR UI systems based on headset controllers (joystick and laser pointer) to assess the efficiency and spontaneity of the former method.

ACKNOWLEDGMENTS

This work was supported by Tallinn University of Technology, University of Limerick, Technological University of the Shannon: Midlands, Midwest, Estonian Research Council, Marie Sklodowska-Curie grant, European Regional Development Fund and Science Foundation Ireland.

REFERENCES

[1] [n. d.]. ROS-Industrial official website. https://rosindustrial.org/ https://rosindustrial.org/.
[2] [n. d.]. Unity Robotics Hub. https://github.com/Unity-Technologies/Unity-Robotics-Hub https://github.com/Unity-Technologies/Unity-Robotics-Hub.
[3] 2023. Accelerated, containerized application development. https://www.docker.com/ https://www.docker.com/.
[4] Thomas Bergs, Sascha Gierlings, Thomas Auerbach, Andreas Klink, Daniel Schraknepper, and Thorsten Augspurger. 2021. The concept of digital twin and digital shadow in manufacturing. *Procedia CIRP* 101 (2021), 81–84.
[5] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
[6] Sachin Chitta. 2016. MoveIt!: An Introduction. In *Studies in Computational Intelligence*. Springer International Publishing, 3–27. https://doi.org/10.1007/978-3-319-26054-9_1
[7] Liang Gong, Åsa Fast-Berglund, and Björn Johansson. 2021. A framework for extended reality system development in manufacturing. *IEEE Access* 9 (2021), 24796–24813.
[8] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.
[9] Rebecca Hetrick, Nicholas Amerson, Boyoung Kim, Eric Rosen, Ewart J. de Visser, and Elizabeth Phillips. 2020. Comparing Virtual Reality Interfaces for the Teleoperation of Robots. In *2020 Systems and Information Engineering Design Symposium (SIEDS)*. 1–7. https://doi.org/10.1109/SIEDS49339.2020.9106630
[10] Kenichi Komatsu, Paula Alavesa, Antti Pauanne, Tuomo Hänninen, Olli Liinamaa, and Ari Pouttu. 2022. Leveraging 5G in cyber-physical system for low-cost robotic telepresence. In *2022 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*. IEEE, 399–404.
[11] Vladimir Kuts, Jeremy A Marvel, Murat Aksu, Simone L Pizzagalli, Martinš Sarkans, Yevhen Bondarenko, and Tauno Otto. 2022. Digital Twin as Industrial Robots Manipulation Validation Tool. *Robotics* 11, 5 (2022), 113.
[12] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and Evaluation of a User Experience Questionnaire. *USAB 2008* 5298, 63–76. https://doi.org/10.1007/978-3-540-89350-9_6
[13] Donghyeon Lee and Young Soo Park. 2018. Implementation of Augmented Teleoperation System Based on Robot Operating System (ROS). In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 5497–5502. https://doi.org/10.1109/IROS.2018.8594482
[14] Praveen Kumar Reddy Maddikunta, Quoc-Viet Pham, B Prabadevi, Natarajan Deepa, Kapal Dev, Thippa Reddy Gadekallu, Rukhsana Ruby, and Madhusanka Liyanage. 2022. Industry 5.0: A survey on enabling technologies and potential applications. *Journal of Industrial Information Integration* 26 (2022), 100257.
[15] Microsoft. [n. d.]. Microsoft/MixedRealityToolkit-Unity: Mixed reality toolkit (MRTK). https://github.com/microsoft/MixedRealityToolkit-Unity https://github.com/microsoft/MixedRealityToolkit-Unity.
[16] László Monostori. 2014. Cyber-physical production systems: Roots, expectations and R&D challenges. *Procedia Cirp* 17 (2014), 9–13.
[17] Dimitris Mourtzis, John Angelopoulos, and Nikos Panopoulos. 2022. Operator 5.0: A survey on enabling technologies and a framework for digital manufacturing based on extended reality. *Journal of Machine Engineering* 22 (2022).
[18] Abdeldjallil Naceri, Dario Mazzanti, Joao Bimbo, Domenico Prattichizzo, Darwin G. Caldwell, Leonardo S. Mattos, and Nikhil Deshpande. 2019. Towards a Virtual Reality Interface for Remote Robotic Teleoperation. In *2019 19th International Conference on Advanced Robotics (ICAR)*. 284–289. https://doi.org/10.1109/ICAR46387.2019.8981649
[19] Cristina Nuzzi, Stefano Ghidini, Roberto Pagani, Simone Pasinetti, Gabriele Coffetti, and Giovanna Sansoni. 2020. Hands-Free: a robot augmented reality teleoperation system. In *2020 17th International Conference on Ubiquitous Robots (UR)*. 617–624. https://doi.org/10.1109/UR49135.2020.9144841
[20] Mikhail Ostanin, Stanislav Zaitsev, Adelia Sabirova, and Alexandr Klimchik. 2022. Interactive Industrial Robot Programming based on Mixed Reality and Full Hand Tracking. *IFAC-PapersOnLine* 55, 10 (2022), 2791–2796. https://doi.org/10.1016/j.ifacol.2022.10.153 10th IFAC Conference on Manufacturing Modelling, Management and Control MIM 2022.
[21] Marco Piccinelli, Andrea Gagliardo, Umberto Castellani, and Riccardo Muradore. 2021. Trajectory planning using Mixed Reality: an experimental validation.

In *2021 20th International Conference on Advanced Robotics (ICAR)*. 982–987. https://doi.org/10.1109/ICAR53236.2021.9659448

[22] Chung Xue Er (Shamaine), Yuansong Qiao, Vladimir Kuts, John Henry, Ken McNevin, and Niall Murray. 2022. Teleoperation of the Industrial Robot: Augmented Reality Application. In *Proceedings of the 13th ACM Multimedia Systems Conference* (Athlone, Ireland) *(MMSys '22)*. Association for Computing Machinery, New York, NY, USA, 299–303. https://doi.org/10.1145/3524273.3532901

[23] J Ernesto Solanes, Adolfo Muñoz, Luis Gracia, Ana Martí, Vicent Girbés-Juan, and Josep Tornero. 2020. Teleoperation of industrial robot manipulators based on augmented reality. *The International Journal of Advanced Manufacturing Technology* 111 (2020), 1077–1097.

[24] Theodoros Togias, Christos Gkournelos, Panagiotis Angelakis, George Michalos, and Sotiris Makris. 2021. Virtual reality environment for industrial robot control and path design. *Procedia CIRP* 100 (2021), 133–138. https://doi.org/10.1016/j.procir.2021.05.021 31st CIRP Design Conference 2021 (CIRP Design 2021).

[25] Baicun Wang, Pai Zheng, Yue Yin, Albert Shih, and Lihui Wang. 2022. Toward human-centric smart manufacturing: A human-cyber-physical systems (HCPS) perspective. *Journal of Manufacturing Systems* 63 (2022), 471–490.

[26] David Whitney, Eric Rosen, Daniel Ullman, Elizabeth Phillips, and Stefanie Tellex. 2018. ROS Reality: A Virtual Reality Framework Using Consumer-Grade Hardware for ROS-Enabled Robots. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 1–9. https://doi.org/10.1109/IROS.2018.8593513

[27] Huitaek Yun and Martin B.G. Jun. 2022. Immersive and interactive cyber-physical system (I2CPS) and virtual reality interface for human involved robotic manufacturing. *Journal of Manufacturing Systems* 62 (2022), 234–248. https://doi.org/10.1016/j.jmsy.2021.11.018

V publication

V

S.L. Pizzagalli, Y. Bondarenko, B. C. Baykara, A. Niidas, V. Kuts, M. Kerm, and T. Otto. Forestry crane immersive user interface for control and teleoperation. In *ASME International Mechanical Engineering Congress and Exposition*, volume 86649, page V02BT02A021. American Society of Mechanical Engineers, 2022

Proceedings of the ASME 2022
International Mechanical Engineering Congress and Exposition
IMECE2022
October 30–November 3, 2022, Columbus, OH

IMECE2022-94975

FORESTRY CRANE IMMERSIVE USER INTERFACE FOR CONTROL AND TELEOPERATION

Simone Luca Pizzagalli[1,*], Yevhen Bondarenko[1], Baris Cem Baykara[1], Alar Niidas[1], Vladimir Kuts[1], Margus Kerm[2], Tauno Otto[1]

[1]Tallinn University of Technology, Department of Mechanical and Industrial Engineering, Tallinn, Harjumaa, Estonia
[2]BFM, Arkna, Rakvere vald, Estonia

ABSTRACT

Timber industry is one of the most relevant economic sectors in Estonia. Automatization of forestry management and harvesting processes optimization are realities also in this specific domain. As much as in other industrial fields adopting the Industry 4.0 paradigm and core technologies, forestry management, log harvesting and the wood processing industry make use of state-of-the-art sensors, Digital Twins and advanced interfaces for the operators. The latter include Extended Reality solutions and remote-control making use of immersive head mounted displays (HMD). This works presents an innovative system for hydraulic forestry crane teleoperation making use of HMD and wide-angle camera stream. The system hardware is installed locally while the software, integrated in Unity, supports the operator in using the crane's native joysticks and controller for the log loading operations. Additional virtual user interface and controls are included in the immersive view and accessible through the same controls and joysticks.

Keywords: Forestry Crane, Teleoperation, Immersive User Interface

1. INTRODUCTION

Estonian timber industry has shown continuous growth in the past years and is one of the most important GDP sectors with a large internal market and considerable export rate [1]. This sector involves many different satellite companies and enterprises active along the production and processing of raw materials, including log harvesting and transportation equipment production. Many digital tools are already employed in this area with software solutions for timber measurement, survey plan or forwarder truck loading [2]. The whole wood industry supply chain seems to be going through a revolution related to the Industry 4.0 technologies and paradigms. This includes heavy use of sensorization, digitalization of the forest environment, harvesting operation and logistics [3]. XR industries are expecting an increase in their sales and turnover in the coming years due to the acceleration of digital technologies adoption aimed at production optimization, cost reduction and an increasing need for remote work [4]. Extended reality (XR) visualization and interaction tools allow for increasing operator efficiency, decreasing task execution time, and the improvement of safety on the workspace [5] especially in scenarios where the operator is in close contact with the machine [6]. Augmented and Virtual Reality (AR/VR) applications have been widely adopted in different domains, including logistics, manufacturing and constructions, as they facilitate and enhance user interaction, system assessment, teleoperation, and training of new personnel [7]. Examples of similar systems can be found in the forestry and timber industry with limitations on the actual operator everyday use, system portability and architecture flexibility. Our work presents an immersive visualization and user interface (UI) for hydraulic forestry crane control, making use of wide-angle cameras, native controller and joysticks, and immersive head mounted display (HMD). The application allows the operator to control the crane remotely, and from the truck cabin without having to move out of the vehicle. The dedicated immersive UI allows the control and visualization of the main crane and cameras visualization parameters. The scalable system aims at the support of a more efficient truck loading task completion and production costs reduction for the trailer producer with an easier installation on different loader crane models.

1.1 Related works

With the increase performance of hardware technology, device portability, bandwidth speed and accessibility, the use of 360 degrees video streaming on HMDs is becoming more and more popular. This allows for an hands free, realistic and immersive experience which have applications in different fields including medicine, entertainment and building construction [8]. Despite the technical challenges of hardware integration and software development, analogous systems seems to offer a valuable solution for teleoperation and remote control. The development of XR based remote control, path planning and programming methods

*Corresponding author: simone.pizzagalli@taltech.ee

for industrial robotic arms is not new [9]. Many crane producer provide simulators for testing, demo purposes and operator coaching. Likewise, recent studies show and increased interest in the development of XR based systems for crane operation simulation, planning and training. The work by [10] proposes a VR based interactive evaluation tool for crane lift operations. The research focuses on cranes used in construction sites where the handling of large objects might be difficult due to limited clearance with other obstacles and structures, or the heavy weight of the transported objects. The system is further developed in [11] providing advanced path planning simulations, facilitating evaluation of safety, training procedures and enhancing the spatial awareness of the operator through the use of HMD visualization. The study by [12] provide similar applications in the same industrial context. The use of immersive VR solutions and HMDs provides safe and repeatable test beds for operator training, speeding up the process and thus greatly reducing the overall costs of the operations. Another critical topic in crane control and manipulation, and which is usually entrusted to the operator, with increased workload and task execution time, is pose estimation and crane boom tracking. Palonen et al. [13] propose an augmented reality system for the support of hydraulic forestry crane operations under difficult visibility conditions. The solution encompasses sensor-enabled synchronization of the crane boom and the tracking of the surrounding environment. Additional graphic guidance is overlapped to the operator camera view through an HMD enabling the real-time visualization of the crane position. Posture estimation and environmental tracking are also tackled in the work by [14] which provides a lightweight 2D scanner-based estimator for the boom tip position and posture in space. The digitalization of crane control and operations requires advanced levels of spatial awareness together with the possibility of setting goal positions and path planning target for semi or fully automatized procedures. Westeberg et al. [15] introduce a virtual reality environment for forestry crane forwarder control and path planning. A twin of the real crane is connected through Local Area Network (LAN) to the real test-bench. The system allows to set the final target or way-points positions of the crane by the means of mouse or joystick input. Motion planning and collision avoidance algorithms are used to calculate the required path for the crane. Additionally, the same authors [16] present an extensive overview of remote control requirements related to different use case scenarios and propose a structured light camera-based environmentally aware system enabling the crane to identify and avoid local obstacles. Miadlicki et al. [17] present a system for real-time monitoring of loader crane site operations making use of a LIDAR scanner and reporting positive results on the system assessment. The study by [18] uses virtual reality to prototype and assess an innovative assortment grapple aimed at optimizing the forwarder loading process. The system utilizes standard crane joysticks controls and Oculus HMD for the visualization for immersive virtual environment test scenario which is eventually used to assess the grapple prototype.

2. MATERIALS AND METHODS

The crane control and manipulation system we present aims at modularity and flexibility. The visualization of the crane grapple operations and log trailer happens through a wide-angle camera installed on the main crane boom. Two additional cameras allow the operator to verify the extension and stability of trailer's legs. The video stream is visualized through Oculus Quest 2 HMD. A dedicated UI is designed using Unity game engine and integrated with the video stream from the cameras, allowing the operator to have an overview of the crane joints' rotation and boom extension. Moreover, the interface allows to control the basic parameters associated both with the crane hydraulic system, visualization and interaction tools (e.g. camera vignette, image exposure). The control of the crane and the navigation within the UI happens through the native joysticks and controller. This allows the user to seamlessly switch from crane joints and boom handling to UI interaction and menu navigation. Both the immersive video stream, joystick controls action mapping and connection with the crane's native controller are implemented as abstract logical layers in the Unity-based application. This solution allows for a modular system architecture enabling the integration of different joysticks and cameras and the installation of the system on different trailer and crane models. The goal is providing an efficient way to replace visualization and control hardware in the future and integrating the system on different crane models without major software modifications. Moreover the computing hardware is positioned in the truck's compartment and powered by truck battery and engine providing a completely autonomous solution for the company and the operator. The following sections explain the architecture of the system and the functioning of each component in greater detail.

2.1 Architecture of the System

The system was designed with modularity in mind, so that it can be adapted to use different camera and joystick controller models with minimal modifications to the core software application. The Unity application (further referred to as "visualizer") is acting as a visualization and control hub, while the communication with the cameras and crane controller are handles by intermediate Python software modules connected to Unity using sockets. Having the modules for communication with the crane implemented in Python allows for easy modification may the hardware requirements change, without requiring changes to the main Unity app. Now, let us explore each aspect of the system in better detail.

2.2 Hardware components

The system hardware comprises three cameras, of which one main, visualizing the boom and grapple operations, and two secondary dedicated to the crane legs extension visualization before and after loading tasks. The main camera is a Teledyne Flir Oryx 10GigE with 12.3 MP resolution [19]. The secondary cameras are two Teledyne Flir Blackfly S GigE with 1.6 MP resolution each [20]. Visualization of the camera streaming happens through Oculus Quest 2 HMD. This headset has improved hardware and visualization quality and allows the operator to work wirelessly and on battery power when necessary. The proposed software runs on a self assembled Desktop PC mounting an ASRock X570 Taichi motherboard, AMD Ryzen 7 5800X CPU, a GeForce RTX 3080 VENTUS graphics card and 64 GB of ram. The system

FIGURE 1: CRANE TRAILER

test-bench is a BFM forestry crane model BMF650 [21]. The crane is installed on a custom trailer, model BMF8T1E, supplied by the same company (see Figure 1). The controllers and joysticks consist of a Parker [22] IQAN MC43 master controller, two IQAN LC5 joysticks and IQAN MD3 monitor interface, also form Parker, and shown in Figure 2.

FIGURE 2: THE CRANE JOYSTICKS

2.3 Camera streaming and visualization system

Connection to the cameras and reading of their video streams is handled by a separate Python application, as can be seen in the software architecture diagram shown in Figure 3. The decision to create an intermediate Python-based module instead of reading camera stream directly in Unity was made to achieve better flexibility of the final system: many camera producers provide Software Developments Kits (SDKs) for their cameras in Python, while SDKs implemented in C# language are not as common. Thus, using Python for the camera module allows to easily modify the code to support a different camera model by switching the used SDK.

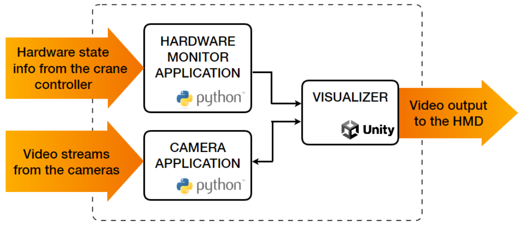

FIGURE 3: SOFTWARE ARCHITECTURE OF THE SYSTEM

The Python camera module is packaged into a server application which handles the connection to the physical camera and re-streams the received video frames to the specified port using TCP protocol. Visualizer connects to the corresponding port and receives the frames through a TCP socket. Such approach allows to swap Python camera modules depending on the used camera model, while keeping the video stream interface the same on Unity side. In order to preserve bandwidth and reduce latency, the frames can be optionally sent in Bayer format (thus reducing the amount of required color channels from 3 to 1) and automatically debayered in Unity to receive full-color images. The received video stream is then applied as a texture to a 3D panel in the visualizer, which displays the view from the camera to the crane's operator.

The camera server also provides an abstract interface for setting camera parameters (such as exposure) through a TCP socket. Combined with the post-processing of received images on Unity side, this allows the operator to apply optimal image settings based on the working conditions using the UI right inside the visualizer.

2.4 Crane controller connection

The crane hardware is controlled by the means of joysticks panel connected to an industrial controller, which handles the immediate control logic and monitors the state of the hardware components. The visualizer application is not integrated into the control loop for safety reasons, and serves purely as a visualization interface for cameras and joystick positions. The data about the state of hardware components (e.g. joysticks) is read from the controller and streamed to Unity by an intermediate Python module, as can be seen in Figure 3. Such design choice was made for the sake of flexibility, similar to the Python camera module.

The used controller (MC43) supports both Generic CAN and SAE J1939 protocols. It has 5 available CAN busses. The scaled percentile values of the analog voltage read-outs have been exposed to an available CAN bus on the MC43 so that a CAN/USB or CAN/ETH adapter could pick up the CAN frames and forward it to a PC. The crane control flow and connection diagram is shown in Figure 4.

In the current implementation, Python hardware monitor module is connecting to the crane controller through a CAN bus and listening to the machine's state by reading CAN frames. The data is then forwarded to a TCP socket, to which the visualizer can connect. As with the camera module, hardware monitor module can be exchanged without breaking the interface on Unity side (e.g., a module using protocol other than CAN to communicate with the controller can be used).

It must be noted that in the default operation mode, the joysticks are mapped to directly control the motion of the crane. Our implementation also allows to halt the control of the crane by pressing a specific button combination on the joysticks. This secondary mode is used to navigate the UI inside the visualizer, e.g. when adjusting the camera image settings. This allows the operator to safely adjust the preferences of the system without risking to move the crane in the process.

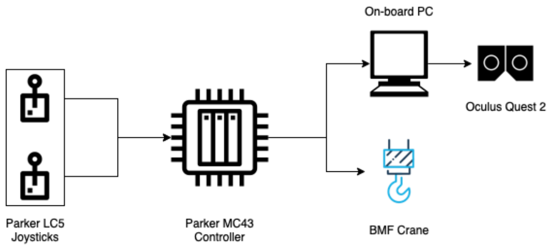

FIGURE 4: CRANE CONTROL FLOW DIAGRAM

2.5 User Interface

The immersive UI integrates some of the functions available from the native crane display and interface provided by the producing company. These functions have been greatly reduced as the project is in an initial stage of development, therefore not all native settings are included in the virtual UI. The main interface section displays real-time information about the crane boom extension, grapple opening, and crane rotation and slewing positions. The layout of this section is shown in Figure 5.

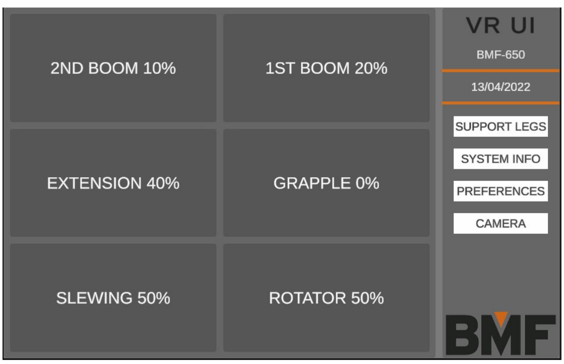

FIGURE 5: UI MAIN SECTION AND CONTROLS

Other interface pages are dedicated respectively to the support leg extension modification, system preferences (e.g., language, date and time) an overview of the system information including modules connections and logs, and camera adjustments (e.g., vignette, exposure and magnification). The operator can shift from crane control to UI navigation by pressing two of the joystick buttons simultaneously. The navigation on the UI elements happens through joystick selection and button press input on either the right or left joystick. The controls are routed and mapped on the existing joysticks through Unity input mapping. Once the UI navigation is selected the crane control system is paused and the operator is able to navigate the interface components. The UI is always visible in overlay view on the camera stream with the above mentioned crane joint information.

2.6 System testing and preliminary results

The application has been successfully tested both in the lab and on the crane provided for the implementation of the system. Preliminary tests has been focusing on camera streaming, crane input controls and interface navigation. The multi camera streaming works correctly and provides the view to the main boom, supporting the operator in the manipulation of the grapple, and the secondary cameras view, dedicated to the trailer legs. The CAN connection was also successfully established allowing for direct input from the physical joysticks to the virtual one and providing effective crane control and immersive UI navigation and interaction. The virtual joysticks, user interface, camera stream visualization in their test setup are shown in Figure 6.

FIGURE 6: CONTROL MODE UI OF THE VISUALIZER

3. CONCLUSION AND FUTURE WORKS

The presented application provides a reliable and scalable solution for forestry crane control, enabling site or remote teleoperation of log truck loading. The architecture allows the provided tools to be easily adjusted to different cranes and hardware components with high level of flexibility and modularity. Unity offers, in this regard, various benefits in terms of HDMs, sensors and other hardware integration possibilities. The use of wide angle cameras allows improved visualization of the working area together with quick access to the main crane controls through immersive UI. Both crane movements and UI navigation and selection are accessed by native crane joysticks and controller allowing more efficient and faster operator learning rate. Furthermore the system allows for a more comfortable and easier operation for the user, especially during harsh weather conditions. The solution presents some limitations especially regarding test and assessment in real world scenario and integration of an extended set of interface functionalities. Further works will finalize the setup on the crane workbench and test the proposed systems with operators in a real world scenario. We expect to compare our solution with the traditional crane operation procedures in controlled experimental sessions. Extensive assessment will be taking into consideration UI design, UI interactions and navigation, camera visualization quality, and crane operation task load and efficiency. The interface will be extended with additional functionalities and diagnostics for the hydraulic system, including further controls and camera adjustments for best image quality visualization. Advanced machine vision safety methods, based on the existing camera setup, could be included to detect obstacles and people in the working area and prevent damages or injuries during the loading operations.

ACKNOWLEDGMENTS

The current research was funded by Tallinn University of Technology R&D applied project for Lisako OU - LEMEE21089 "Product development project: Application of virtual reality for the crane control" and conducted using the Smart Industry Centre (SmartIC) core facility funded by the Estonian Research Council grant TT2.

REFERENCES

[1] Misc, A. "Forestry." Keskkonnaministeerium (2021). Accessed March 27, 2022, URL https://envir.ee/en/water-forest-resources/forestry.

[2] Misc, A. "Digital revolution in the Estonian forestry and wood industry." Estonian Timber (2018). Accessed March 27, 2022, URL https://estoniantimber.ee/best-practices/digital-revolution-in-the-estonian-forestry-and-wood-i/.

[3] Müller, Fabian, Jaeger, Dirk and Hanewinkel, Marc. "Digitization in wood supply–A review on how Industry 4.0 will change the forest value chain." *Computers and Electronics in Agriculture* Vol. 162 (2019): pp. 206–218.

[4] Misc, A. "XR and its potential for Europe." XR Association (2021). Accessed March 8, 2022, URL https://xra.org/research/xr-and-its-potential-for-europe/.

[5] Malik, Ali Ahmad, Masood, Tariq and Bilberg, Arne. "Virtual reality in manufacturing: immersive and collaborative artificial-reality in design of human-robot workspace." *International Journal of Computer Integrated Manufacturing* Vol. 33 No. 1 (2020): pp. 22–37. DOI 10.1080/0951192X.2019.1690685.

[6] Dianatfar, Morteza, Latokartano, Jyrki and Lanz, Minna. "Review on existing VR/AR solutions in human–robot collaboration." *Procedia CIRP* Vol. 97 (2021): pp. 407–411.

[7] Doolani, Sanika, Wessels, Callen, Kanal, Varun, Sevastopoulos, Christos, Jaiswal, Ashish, Nambiappan, Harish and Makedon, Fillia. "A Review of Extended Reality (XR) Technologies for Manufacturing Training." *Technologies* Vol. 8 No. 4. DOI 10.3390/technologies8040077. URL https://www.mdpi.com/2227-7080/8/4/77.

[8] Shafi, Rabia, Shuai, Wan and Younus, Muhammad Usman. "360-degree video streaming: A survey of the state of the art." *Symmetry* Vol. 12 No. 9 (2020): p. 1491.

[9] Kuts, Vladimir, Otto, Tauno, Tähemaa, Toivo and Bondarenko, Yevhen. "Digital twin based synchronised control and simulation of the industrial robotic cell using virtual reality." *Journal of Machine Engineering* Vol. 19.

[10] Kayhani, Navid, Taghaddos, Hosein, Noghabaee, Mojtaba et al. "Utilization of virtual reality visualizations on heavy mobile crane planning for modular construction." *arXiv preprint arXiv:1901.06248* .

[11] Pooladvand, Shiva, Taghaddos, Hosein, Eslami, Arash, Nekouvaght Tak, Ala and Hermann, Ulrich. "Evaluating Mobile Crane Lift Operations Using an Interactive Virtual Reality System." *Journal of Construction Engineering and Management* Vol. 147 No. 11 (2021): p. 04021154.

[12] Patrão, Bruno and Menezes, Paulo. "A virtual reality system for training operators." *International Journal of Online Engineering (iJOE)* Vol. 9 No. S8 (2013): pp. 53–55.

[13] Palonen, Tuomo, Hyyti, Heikki and Visala, Arto. "Augmented reality in forest machine cabin." *IFAC-PapersOnLine* Vol. 50 No. 1 (2017): pp. 5410–5417.

[14] Hyyti, Heikki, Lehtola, Ville V and Visala, Arto. "Forestry crane posture estimation with a two-dimensional laser scanner." *Journal of Field Robotics* Vol. 35 No. 7 (2018): pp. 1025–1049.

[15] Westerberg, Simon, Manchester, Ian R, Mettin, Uwe, La Hera, Pedro and Shiriaev, Anton. "Virtual environment teleoperation of a hydraulic forestry crane." *2008 IEEE International Conference on Robotics and Automation*: pp. 4049–4054. 2008. IEEE.

[16] Westerberg, Simon and Shiriaev, Anton. "Virtual environment-based teleoperation of forestry machines: Designing future interaction methods." *Journal of Human-Robot Interaction* Vol. 2 No. 3 (2013): pp. 84–110.

[17] Miadlicki, Karol, Pajor, Miroslaw and Sakow, Mateusz. "Loader crane working area monitoring system based on LIDAR scanner." *Advances in Manufacturing*. Springer (2018): pp. 465–474.

[18] Manner, Jussi, Berg, Simon, Englund, Martin, Ersson, Back Tomas and Mörk, Anders. "Innovative productivity improvements in forest operations: a comparative study of the Assortment Grapple using a machine simulator." *Journal of Forest Science* Vol. 66 No. 11 (2020): pp. 443–451.

[19] "Camera: Oryx 10GigE ORX-10GS-123S6C-C." URL https://www.flir.eu/products/oryx-10gige/?model=ORX-10GS-123S6C-C.

[20] "Camera: Blackfly S GigE BFS-PGE-16S2C-CS." URL https://www.flir.com/products/blackfly-s-gige/?model=BFS-PGE-16S2C-CS.

[21] Misc, A. "BMF Forestry Trailers & Cranes." BMF (2021). Accessed April 05, 2022, URL https://www.bmf.ee/products/cranes/bmf-650.

[22] Misc, A. "Parker Engineering Your Success Motion Control Technology." Parker (2022). Accessed April 05, 2022, URL https://www.parker.com.

VI publication

VI

K. Mahmood, **S.L. Pizzagalli**, T. Otto, and I. Symotiuk. Development of an AR-based application for assembly assistance and servicing. In *Procedia CIRP, 34th CIRP design Conference, 3-5 June 2024, Cranfield University, UK*. CIRP, 2024

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 00 (2023) 000–000

www.elsevier.com/locate/procedia

34th CIRP Design Conference

Development of an AR-based application for assembly assistance and servicing

Kashif Mahmood[a]*, Simone Luca Pizzagalli, Tauno Otto, Ivan Symotiuk

[a]Tallinn University of Technology, Ehitajate tee 5, Tallinn 19086, Estonia

* Corresponding author. Tel.: +372 6203253; *E-mail address:* kashif.mahmood@taltech.ee

Abstract

Recent advancements in the field of digital manufacturing, especially the adoption of fast connectivity through 5G, Digital Twins and Extended Reality (XR) in manufacturing, offer new possibilities for innovative and effective design of production workflows. Augmented Reality (AR) can assist in speeding up the assembly and maintenance processes by facilitating the operators to perform these processes without dealing with detailed paper manuals. AR-based interactive user interfaces can support operators to be more productive by visualizing certain product and component 3D models dynamically, in addition to assembly steps along the corresponding manufacturing process. Likewise, AR-based applications can facilitate the setup and maintenance service of a machine by providing advanced machine visualization and digitalized information. This paper presents a conceptual model and case-based demo applications adopting AR technology for the maintenance of an industrial machine and supporting operators during the assembly process of a specific product. The goal is to improve productivity by reducing the processing time and minimize the operator training time. Moreover, the proposed AR application is integrated with the non-conformance reporting feature which helps to address the quality related issues quickly and efficiently, leading to a reduction in the number of nonconformities. Two case studies demonstrate the relevance of the proposed conceptual model and testing of the applications.

Keywords: Augmented reality, digital work instruction, user interface design, assembly and maintenance

1. Introduction

Since the introduction of the concept of Industry4.0 (I4.0) [1], digitalization has become one of driving forces in manufacturing process and product design. The hyperconnected modern industrial world has its roots in the process of integration of physical assets with the digital domain [2]. The combination of physical and digital realms has been eased and fostered by the continuous development and improvement of related technologies such as sensors, high speed connectivity, automation, Digital Twins (DT) and Human Machine Interfaces (HMI). I4.0 has been developing from the mere automatization to a scenario that demands operators back in the process to fulfil the higher requirements for a resilient production and highly customized products. An increasingly human centric approach drives the new digital tools development by revolving around operators and machines interactions rather than data only. This is known as the Industry5.0 (I5.0) revolution. Edge computing, advanced DTs, high speed mobile connectivity (5G) and HMI methods through extended reality (XR) user interfaces (UIs) are some of the main technologies at the base of this transition [3] supporting operators' performance and wellbeing together with higher production efficiency and flexibility. The match of DTs and XR UIs technologies seems to answer some of the new I5.0 requirements by including the operators in the loop and providing ways to visualize and interact with both digital contents and physical machines. These technologies have already been implemented and tested in operator training, robot programming, remote maintenance, assembly support use cases [4] replacing old practices and analogue information handling. Regardless, the adoption rate, and perceived advantages, for Smaller and Medium Enterprises (SMEs) even in case of I4.0, are slowed down by manifold challenges, including lack of resources (time and financial), lack of support and training in the use of the technologies, limited awareness and potentials of the available tools [5].

This work presents the design flow and implementation of DT based AR UIs for two different SMEs in the Estonian timber industry. The first use case application goal is to improve machine maintenance, by reducing the task time and providing an easily understandable access to the procedures for every operator. The second use case aims at supporting operators in the assembly process of wooden house modules. The use case leverages on a local 5G network connectivity and integration of a Product Lifecycle Management (PLM) software for products data, assembly procedures and product 3D models file integration. The application aims at reducing the assembly task time, improving the product redesign flow, fasten the production line changeover while reducing nonconformities and assembly errors.

2. Related Works

Human centricity and resiliency are two of core values for the EU industrial vision for the future [6]. The attention to highly resilient agents, such as humans, is aimed at a fast response to product demand changes and disruption. Actions going in this direction are expected, with a focus on what technology can do to support operators in the production processes and collaboration with machines [7]. As much as technologies keep improving and developing at a fast pace, the game changer is the shift in the overall approach and focus on User Centered Design (UCD) practices [8].

DT - XR systems, and in general advanced HMIs, allow to put the operators back amid data and physical assets, providing an access point to the digital world, supporting product verification routines, teleoperation, maintenance, assembly processes and training [9]. The difficulties in implementing these systems, especially in the case of AR, are manifold [4], use case related, and requiring a comprehensive UCD design framework supporting an effective and efficient deployment in different scenarios.

Simulations and DTs technologies are essential tools in manufacturing, facilitating, if not driving, every stage of the product lifecycle, from design conception to product disposal and materials reuse [10]. These technologies are used in combination with PLM software to manage data and design changes between the physical assets and digital contents at every level of the production process. As already mentioned, high speed connectivity based on 5G is another driving force in the context of I5.0. Fifth generation mobile networks enable large scale IoT implementation, data and video communication, edge computing and the deployment flexible manufacturing systems [11]. While supposedly beneficial for smaller companies in speeding up the digitalization process and smart manufacturing practices [12], the adoption of these networks, especially in SMEs, is still in its infancy and need more use cases and testbeds to balance costs and benefits and understand the ways forward [13].

XR and more specifically AR, show a considerable increase in applications and examples in many industrial fields [14]. The study described in [15] presents a framework for product design development in a collaborative cloud-based environment. The scope is the transfer of competencies between academia to industry. The specific use case supports desktop based, AR and handheld devices. Another collaborative application adopting AR visualization and advanced interactions aimed at operator training and assembly guidance support is presented in [16]. The system supports the simultaneous visualization and interaction of different operators working on large components in the context of ship building. The system assessment seems to endorse the usefulness of the solution while many problems are still open, especially related to hardware and software compatibility. An assembly instruction interactive projected augmented reality application is described in [17]. A simple use case based on the assembly of toy blocks is proposed and assessed for efficiency, number of errors and task load. Regardless of the limitations of the system, result confirm the reduced effort, and number of errors during the assembly while adopting AR. The study in [18] proposes a combination of Head Mounted Display (HMD) and human simulated based guidance for worker assistance in an assembly step. The system is validated and assessed during an experimental session comparing the scenarios with old style pictorial assembly guidance showing high usability and task efficiency for the XR solutions. Another comparative study supporting the benefits of AR systems for training and assembly support is presented in [19] The study also argues about the need for adaptivity of technological systems for each operator while fostering creative and problem-solving skills at the same time. However, there is a lack of systematic harmonized and hardware agnostic approaches implementing AR-based solutions to assist and support operators for assembly and maintenance tasks. This study provides a holistic development of AR application leveraging 5G technology which not only integrates UCD in the application design and development but also validates it through industrial use cases.

3. Context of Use Cases

The proposed solutions grounds on two different use cases in the scope of timber industry and favoring the acceleration of the digitalization processes for SMEs through 5G connectivity networks. The first use case (UC1) aims at aiding maintenance and servicing of machinery involved in timber products handling and packing. The second use case (UC2) focuses on a support interface for wooden house modules assembly procedures. The UCs relate to two different Estonian SMEs active in the timber industry on local and European level.

3.1. Design Process and Application Requirements

As already mentioned, a UCD process is adopted for the specific applications. This involves an iterative development combining requirement definition and assessment steps with UI development and refinement of the interface functions. The requirements definition for the two use cases are based on an initial survey with the company's management and operators team leaders. The data structuring and maintenance scheduling of the UC1 is in the development phase. The identified needs for this UC are structured as follows:

- Augmented reality instructions and content visualization.
- Image target/QR code-based localization on the specific machine.
- Simple UI guidance through the servicing tasks.

UC2 development is integrating digitalized 3D information and procedures. The company already provided feedback on the application to optimize the integration of existing methods and IT solutions. Regardless, the initial requirements for the XR application presented in this study are the following:

- System login through operator QR code.
- Dashboard to select the desired product or component (also accessible through product related barcode)
- Dynamic (animated) visualization of 3D component models throughout the assembly process.
- AR visualization of the abovementioned models.
- Integration of assembly guidance information.
- Integration of nonconformity reporting system with text audio and photo-based feedback
- Application should be running on different devices and integrated with PLM repository.

3.2. Key Performance Indicators

Key Performance Indicators (KPIs) are the driving force for the performance evaluation and measurement of any system, including the effectiveness of technological solutions and strategic decisions. Their role is also to drive and magnify the improvement of a process [20]. Based on the abovementioned general goals and requirements the proposed applications are intended to demonstrate the benefits of local high-speed connectivity in allowing advanced XR visualization for operator support on the production floor. The following KPIs are taken into consideration while developing the UIs:

- Servicing and faults reporting time.
- Operator training time.
- Process lead time.
- Number of nonconformance and reporting time.

4. System Architecture

Two different systems were proposed in terms of software technologies adopted and related architecture.

4.1. Use case 1

The first application is on a demonstration, interactive mockup level. As the use case target machine is installed in an Estonian third company sawmill facility, the development is happening locally using a laboratory Flexible Manufacturing System (FMS) equipment as test case. The application is installed and running on an android based tablet. All information, at this stage, is stored locally on the device. The interface is developed by using Unity game engine [21] and native UI packages. Vuforia SDK [22] is integrated in Unity allowing for image target recognition and 3D digital content visualization on real world objects. The target images and QR codes allow to position the augmented content/instructions in real space and retrieve the correct information for the maintenance procedure.

4.2. Use Case 2

Following the requirements of the owner company the application is accessible from any device, handheld or head mounted, including the tablets used on the production floor. Database information (PLM) is stored on the company server. Local 5G connectivity provides safe access to the main repository and synchronized data flow from company design department. The plan is to implement 5G hardware and software core setup, however, local secured Wi-Fi connectivity is employed for the preliminary tests.

The application is developed by using of Next.js framework [23], Tailwind.css library [24], for advanced UI elements integration and styling, and Three.js [25], for the integration and visualization of product and components 3D models. The architecture of the system, Fig. 1, consists of two main components: a Next.js application and a browser-based interface app.

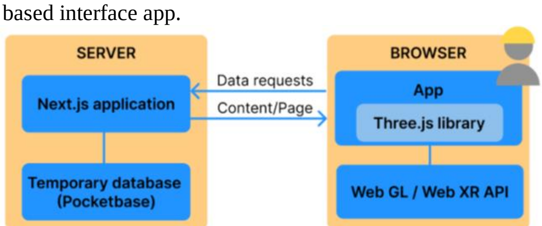

Fig. 1. UC2 system architecture

The first application runs on the local server where a provisional database of instructions and models is stored, in this case – PocketBase [26], and which later should be replaced with PLM repository. The Next application sends the page template to the user's browser where information is displayed as well as 3D models are rendered using device's native hardware. HTTP requests to API endpoints are sent to retrieve assembly information or post nonconformities. The placement of the 3D content in real world by AR visualization happens, in this case, using Three.js XR functionality which utilizes WebXR device API [27]. This is not image target based but uses ground detection functions. This allows to detect horizontal surfaces through the device camera and place the 3D content in space accordingly.

5. User Interface

As in both cases the preferred device for the required tasks and the specific working environments is a tablet, user interactions with the UIs is happened through tap gesture. Pinch gesture is used, specifically in UC2, for model rotation and zoom.

5.1. UI for Use Case 1

The information architecture for the UC1 UI application is shown in Fig. 2.

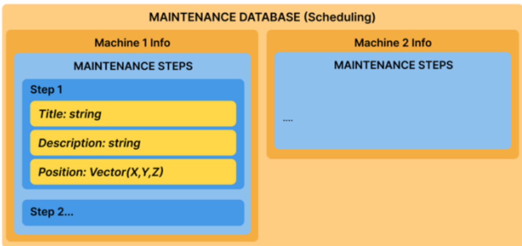

Fig. 2. Information architecture for UC1 UI application.

The information about machine maintenance scheduling is saved locally on the device. Future developments will retrieve the information from the producer company server through specific APIs and update the instructions dynamically. Each machine holds information on the scheduled servicing with related steps and extended descriptions of the actions the operator must perform to fulfill the tasks. These are localized by QR codes and image targets on the real machine and visualized through augmented information cards. An example of AR visualization of the maintenance step card is shown is Fig. 3. The picture in picture image shows the operator holding the tablet and targeting the QR code (bottom left), and the augmented maintenance step card (larger image) visualized on the tablet screen in a relative position to the targeted images.

Fig. 3. Picture in Picture visualization of the maintenance UI application.

5.2. UI for Use Case 2

In UC2 the application has a more complex structure and functional capabilities. Fig. 4 shows the information architecture for the developed software.

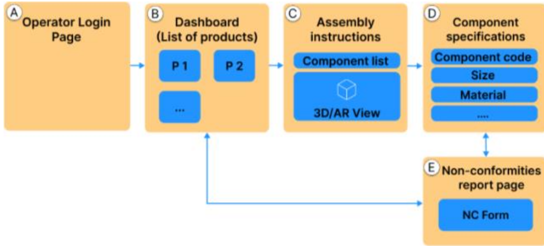

Fig. 4. Information architecture of the assembly support application.

Five main content blocks are envisioned for the application including a login page (A), a project dashboard (B), an assembly instruction block (C), which can contain different type of information depending on the component to be assembled, a component specification block (D), also updated dynamically, and a nonconformity report block (E). Operators can login in the application by scanning their personal employee card QR code. Block B in the diagram constitutes the application dashboard page, where the operator can access different products and individual product components. These are visualized as 3D models previews together with an assembly status bar showing how much of assembly work has already been completed. Block C represents the assembly instructions page, Fig. 5.

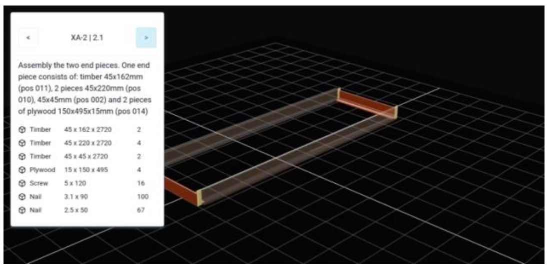

Fig. 5. Assembly step page (tablet screenshot)

This contains assembly steps information for each component part together with a dynamic and animated 3D/AR visualization. At each step the part that needs to be assembled is visualized in the model while assembly instructions are shown as a list to the worker. The 3D model can be navigated through pinch rotate and zoom gestures. Fig. 6 shows AR visualization of one of the components with the assembly instructions.

Fig. 6. Assembly step page with AR visualization.

If the worker needs more information on a specific part, the name of the part can be clicked in the assembly list and open a pop-up window with further details. These include component identification code, measurements, material, and number of pieces. From the part specification pop-up window, it is possible to access the nonconformities report page, block E. This page can be accessed also from the main application dashboard. The functionality allows to report a problem with the currently selected part. The interface allows to select from a list of pre-defined nonconformities, upload images, text, and audio descriptions of the observed issue.

6. Preliminary Assessment

The UC2 application has been preliminary tested in the owner company. A group of 6 subject has been recruited for the assessment including different professional profiles, from operators on the assembly line to management. The experimental protocol consists of the following steps. After a preliminary briefing, users were asked to sign an informed consent and perform two tasks. The first was an intuitive navigation of the application; the second consists in reporting two nonconformities of specific parts involved in the assembly procedure of two different product components. After concluding the tasks, the subjects completed three different questionnaires.

6.1. Assessment Metrics

The assessment of the system focuses on the evaluation of user experience and usability based on System Usability Scale (SUS) [28] and User Experience Questionnaire (UEQ)[29]. A custom questionnaire was used to gather generic demographics data and specific feedback on UI elements visibility and accessibility. SUS questionnaire comprises ten items rated on a five-point Likert scale, assessing the user's overall perception of system effectiveness, efficiency, and satisfaction. Results are scored on a scale ranging from 0 to 100, with values below 50 corresponding to non-acceptable usability, between 50 and 70 to acceptable, between 70 and 85 to good usability with values above this corresponding to excellent. The UEQ short version consists of 8 items that capture hedonic and pragmatic aspects of the interaction with the application, rated on a 7-point Likert scale, each with both positive and negative terms. Scores range from -3 to 3, indicating the spectrum from very poor to very good user experience (UX). Ratings between -0.8 and 0.8 are neutral or average, while scores above 0.8 signify a positive UX assessment, and those below -0.8 indicate a negative evaluation. The custom questionnaire items were evaluated on a five-points Likert scale with a negative and a positive term for each statement. A speak aloud protocol was also adopted during the test allowing users to give comments on their experience. The comments have been recorded as a base for further improvements and partially reported in the results section.

6.2. Results

The sample comprised two factory workers, one pre-production employee, one 2D department manager, one employee in the IT department and one in management. The average age of the subject sample is 35 years old (SD of ±7).

The SUS questionnaire results show good overall usability of the application with an average of 72.9 (SD ±9.7) corresponding to an acceptable usability level.

Results of the UEQ show a very good response for pragmatic qualities of the application (2.54) and an overall UX scoring of 2.1. These results are followed by hedonic quality with a result of 1,67 which is in any case more than good.

The custom questionnaire gave positive results in terms of clarity of the UI element position and functions. The overall navigation, and visualization of the 3D elements were adequate. Results are reported in Table 1.

Table 1. Custom questionnaire results

	Mean	SD
Interface buttons are clearly visible	4,3	0,7
The application controls are clear and accessible	4,3	0,5
The component 3D visualisation is intuitive and clear	4,8	0,4
The navigation of the interface is complicated	3,5	0,9
I could not find the pages I needed to access	2,3	1,6
Functions I needed are easy to find	3,8	0,6

The speak aloud protocol reveals a few problems with the navigation of the application pages especially between the additional documentation materials and the assembly steps. One user suggested to decrease the sensitivity of the navigation of 3D elements and include pinch to zoom functionalities in AR mode. Another subject pointed out that the activation of the AR mode should be more intuitive. Breadcrumbs might be included during the assembly procedure to allow the operator an easier assessment of the state of the work process.

7. Conclusion and Future Work

The proposed applications contributed to providing a dynamic XR visualization and support operators in industry related tasks, such as machine maintenance and product assembly, through the use of local 5G connectivity. The application fits into the course of SMEs digitalization of industrial processes and adoption of I4.0 technologies and I5.0 user centered approach. The two AR based interfaces have been developed adopting a recursive and agile method by fulfilling the requirements and needs of the companies while adapting to the ongoing digitalization process. The UC2 application has been preliminary tested within the owner company with satisfactory results in terms of Usability and User Experience.

Further developments will consider the comments and data gathered during the test for system improvements related to UI development. Subsequently the completion of data repositories, PLM software and installation of the local 5G network, the applications will be integrating dynamic data retrieval through REST APIs. A medium- and long-term assessment plan for the proposed KPIs will take place in the company premises once the abovementioned related systems will be installed and fully operative. These will involve a larger group of operators and team leaders and provide essential data about the overall efficiency and effectiveness of all AR application related technological solutions and proposed interfaces.

Acknowledgements

The research was supported and conducted using the "Secure 5G-Enabled Twin Transition for Europe's TIMBER Industry Sector" funded by European Union's Horizon Europe Research program (5G-TIMBER).

References

[1] H. Kagermann, W. Wahlster, Ten Years of Industrie 4.0, Sci 4 (2022) 26. https://doi.org/10.3390/sci4030026.
[2] L. Monostori, Cyber-physical Production Systems: Roots, Expectations and R&D Challenges, Procedia CIRP 17 (2014) 9–13. https://doi.org/10.1016/j.procir.2014.03.115.
[3] P.K.R. Maddikunta, Q.-V. Pham, P. B, N. Deepa, K. Dev, T.R. Gadekallu, R. Ruby, M. Liyanage, Industry 5.0: A survey on enabling technologies and potential applications, J Ind Inf Integr 26 (2022) 100257. https://doi.org/10.1016/j.jii.2021.100257.
[4] L.F. de Souza Cardoso, F.C.M.Q. Mariano, E.R. Zorzal, A survey of industrial augmented reality, Comput Ind Eng 139 (2020) 106159.
[5] T. Masood, P. Sonntag, Industry 4.0: Adoption challenges and benefits for SMEs, Comput Ind 121 (2020) 103261. https://doi.org/10.1016/j.compind.2020.103261.
[6] E. Commission, D.-G. for Research, Innovation, M. Breque, L. De Nul, A. Petridis, Industry 5.0 – Towards a sustainable, human-centric and resilient European industry, Publications Office of the European Union, 2021. https://doi.org/doi/10.2777/308407.
[7] D. Mourtzis, J. Angelopoulos, N. Panopoulos, Operator 5.0: A Survey on Enabling Technologies and a Framework for Digital Manufacturing Based on Extended Reality, Journal of Machine Engineering 22 (2022) 43–69. https://doi.org/10.36897/jme/147160.
[8] S.L. Pizzagalli, V. Kuts, T. Otto, User-centered design for Human-Robot Collaboration systems, IOP Conf Ser Mater Sci Eng 1140 (2021) 012011. https://doi.org/10.1088/1757-899X/1140/1/012011.
[9] Y. Yin, P. Zheng, C. Li, L. Wang, A state-of-the-art survey on Augmented Reality-assisted Digital Twin for futuristic human-centric industry transformation, Robot Comput Integr Manuf 81 (2023) 102515. https://doi.org/10.1016/j.rcim.2022.102515.
[10] C.K. Lo, C.H. Chen, R.Y. Zhong, A review of digital twin in product design and development, Advanced Engineering Informatics 48 (2021) 101297. https://doi.org/10.1016/j.aei.2021.101297.
[11] M. Attaran, The impact of 5G on the evolution of intelligent automation and industry digitization, J Ambient Intell Humaniz Comput 14 (2023) 5977–5993. https://doi.org/10.1007/s12652-020-02521-x.
[12] M. Olokundun, M.E. Ogbari, H. Falola, A.S. Ibidunni, Leveraging 5G network for digital innovation in small and medium enterprises: a conceptual review, J Innov Entrep 11 (2022) 41. https://doi.org/10.1186/s13731-021-00181-5.
[13] D.S. Fowler, Y.K. Mo, A. Evans, S. Dinh-Van, B. Ahmad, M.D. Higgins, C. Maple, A 5G Automated-Guided Vehicle SME Testbed for Resilient Future Factories, IEEE Open Journal of the Industrial Electronics Society 4 (2023) 242–258. https://doi.org/10.1109/OJIES.2023.3291234.
[14] M. Eswaran, A.K. Gulivindala, A.K. Inkulu, M.V.A. Raju Bahubalendruni, Augmented reality-based guidance in product assembly and maintenance/repair perspective: A state of the art review on challenges and opportunities, Expert Syst Appl 213 (2023) 118983. https://doi.org/10.1016/j.eswa.2022.118983.
[15] D. Mourtzis, V. Siatras, J. Angelopoulos, N. Panopoulos, An Augmented Reality Collaborative Product Design Cloud-Based Platform in the Context of Learning Factory, Procedia Manuf 45 (2020) 546–551. https://doi.org/10.1016/j.promfg.2020.04.076.
[16] A. Vidal-Balea, O. Blanco-Novoa, P. Fraga-Lamas, M. Vilar-Montesinos, T.M. Fernández-Caramés, Creating Collaborative Augmented Reality Experiences for Industry 4.0 Training and Assistance Applications: Performance Evaluation in the Shipyard of the Future, Applied Sciences 10 (2020) 9073. https://doi.org/10.3390/app10249073.
[17] M. Simonetto, M. Peron, G. Fragapane, F. Sgarbossa, Digital assembly assistance system in Industry 4.0 era: a case study with projected augmented reality, in: Advanced Manufacturing and Automation X 10, 2021: pp. 644–651.
[18] E. Lampen, J. Teuber, F. Gaisbauer, T. Bär, T. Pfeiffer, S. Wachsmuth, Combining Simulation and Augmented Reality Methods for Enhanced Worker Assistance in Manual Assembly, Procedia CIRP 81 (2019) 588–593. https://doi.org/10.1016/j.procir.2019.03.160.
[19] M. Moghaddam, N.C. Wilson, A.S. Modestino, K. Jona, S.C. Marsella, Exploring augmented reality for worker assistance versus training, Advanced Engineering Informatics 50 (2021) 101410. https://doi.org/10.1016/j.aei.2021.101410.
[20] K. Mahmood, M. Lanz, V. Toivonen, T. Otto, A Performance Evaluation Concept for Production Systems in an SME Network, in: Procedia CIRP, Elsevier B.V., 2018: pp. 603–608. https://doi.org/10.1016/j.procir.2018.03.182.
[21] Unity Game Engine, (n.d.). https://unity.com/ (accessed November 17, 2023).
[22] Vuforia for Unity, (n.d.). https://developer.vuforia.com/library/getting-started/getting-started-vuforia-engine-unity (accessed November 17, 2023).
[23] Next.JS the REACT framework, (n.d.). https://nextjs.org/ (accessed November 17, 2023).
[24] Tailwind CSS, (n.d.). https://tailwindcss.com/ (accessed November 17, 2023).
[25] Three.js, (n.d.). https://threejs.org/ (accessed November 17, 2023).
[26] Pocketbase, (n.d.). https://pocketbase.io/ (accessed November 17, 2023).
[27] Web XR, (n.d.). https://www.w3.org/TR/webxr/ (accessed November 17, 2023).
[28] J.Brooke, SUS: A "Quick and Dirty" Usability Scale, in: Usability Evaluation In Industry, CRC Press, 1996: pp. 207–212. https://doi.org/10.1201/9781498710411-35.
[29] B. Laugwitz, T. Held, M. Schrepp, Construction and evaluation of a user experience questionnaire, in: Lecture Notes in Computer Science, Springer Verlag, 2008: pp. 63–76. https://doi.org/10.1007/978-3-540-89350-9_6.

Curriculum Vitae

1. Personal data

Name	Simone Luca Pizzagalli
Date and place of birth	1 February 1983, Merate, Italy
Nationality	Italian

2. Contact information

Address	Tallinn University of Technology, School of Engineering Department of Mechanical and Industrial Engineering Ehitajate tee 5, 19086 Tallinn, Estonia
E-mail	simone.pizzagalli@taltech.ee

3. Education

2020–2024	Tallinn University of Technology, School of Engineering, PhD studies
2015–2016	Italian National Research Council, Politecnico di Milano, Design for All: Advanced Human Machine Interfaces
2005–2007	Delft University Of Technology, Faculty of Architecture, Architecture, urban planning and building sciences, MSc, *cum laude*
2008–2011	Politecnico di Milano, Faculty of Architecture, Architecture and urban planning, BSc

4. Language competence

Italian	native
English	fluent
Spanish	basic level

5. Professional employment

2020– . . .	Tallinn University of Technology, School of Engineering, Department of Mechanical and Industrial Engineering, Early stage researcher
2017–2020	Italian National Research Council, STIIMA, Lecco, Italy, Early stage researcher
2018–2018	Soongsil University, Seoul, South Korea, Visiting Professor
2016–2016	Politecnico di Milano, Milan, Italy, Research Internship
2012–2014	Osar n.v., Antwerp, Belgium, Designer, Researcher
2010–2014	Menabo' architecture, Rotterdam, The Netherlands, Designer
2008–2010	De Werff Architectuur, The Hague, The Netherlands, Designer
2007–2007	ZUS, Zones Urbaines Sensibles, Rotterdam, The Netherlands, Internship Designer

6. Voluntary work

2021–2023	NOOR MTÜ, Estonia.
2014–2016	Piccoli Artigiani. Participatory design and planning, interior refurbishment with scrap materials, Italy.
2010–2014	Founder of POST—OFFICE architectural platform in Rotterdam: renovation and activation of an abandoned location in Rotterdam North; participatory design activities and projects; promotion of architectural architectural and artisanal culture, The Netherlands.

7. Honours and awards

- 2023 Best presenter at 2023 International Conference on Computers and Automation
- 2023 Best Demo Award at EUROXR 2023: ROS based XR interface for robot path-planning and teleoperation.
- 2018 H4O Hackathon for Ophthalmology Winner of category CARE / Winner of Fondazione Cottino Price Project proposal for a device aimed at the improvement of care solutions for visually impaired people.
- 2011 Archiprix International 2011, New York First Prize for Best Graduation Project.

8. Field of research

- Human Machine Interaction
- Extended Reality
- User Experience

9. Scientific work
Papers not included in the thesis

1. W. Terkaj, E. Pessot, V. Kuts, Y. Bondarenko, S. L. Pizzagalli, and K. Kleine. A framework for the design and use of virtual labs in digital engineering education. In *AIP Conference Proceedings*, volume 2989-1. AIP Publishing, 2024
2. Y. Bondarenko, S. L. Pizzagalli, V. Kuts, E. Petlenkov, and T. Otto. Digital twin based learning from demonstration system for industrial robots. In *ASME International Mechanical Engineering Congress and Exposition*, volume 87608, page V003T03A081. American Society of Mechanical Engineers, 2023
3. Y. Bondarenko, V. Kuts, S. Pizzagalli, K. Nutonen, N. Murray, and E. O'Connell. Universal xr framework architecture based on open-source xr tools. In *International XR Conference*, pages 87–98. Springer, 2023
4. M. Mondellini, S. Arlati, H. Gapeyeva, K. Lees, I. Märitz, S. L. Pizzagalli, T. Otto, M. Sacco, and A. Teder-Braschinsky. User experience during an immersive virtual reality-based cognitive task: a comparison between estonian and italian older adults with mci. *Sensors*, 22(21):8249, 2022
5. K. Kalda, S.-L. Pizzagalli, R.-M. Soe, R. Sell, and M. Bellone. Language of driving for autonomous vehicles. *Applied Sciences*, 12(11):5406, 2022
6. D. Diachenko, A. Partyshev, S. L. Pizzagalli, Y. Bondarenko, T. Otto, and V. Kuts. Industrial collaborative robot digital twin integration and control using robot operating system. *Journal of Machine Engineering*, 22, 2022
7. V. Kuts, Y. Bondarenko, M. Gavriljuk, A. Paryshev, S. Jegorov, S. Pizzagall, and T. Otto. Digital twin: Universal user interface for online management of the manufacturing system. In *ASME International Mechanical Engineering Congress and Exposition*, volume 85567, page V02BT02A003. American Society of Mechanical Engineers, 2021

8. R. Nossa, M. Porro, O. Picciolini, M. Malosio, S. Pittaccio, M. Valoriani, V. Asperti, F. Clasadonte, L. Oliveto, M. Mondellini, et al. What children with neuromotor disabilities need to play with technological games. *Applied Sciences*, 11(21):9886, 2021

9. L. Ordile, Y. Bondarenko, S. L. Pizzagalli, V. Kuts, and T. Otto. Augmented reality interface for industrial robot control and teleoperation. *nil e H ja K*, 2021

10. M. Mondellini, S. Arlati, S. Pizzagalli, L. Greci, M. Sacco, G. Ferrigno, and M. Sacco. Assessment of the usability of an immersive virtual supermarket for the cognitive rehabilitation of elderly patients: a pilot study on young adults 2018 ieee 6th international conference on serious games and applications for health (segah), 1–8 https://doi. org/10. 1109. *SeGAH*, 84013:13, 2018

11. M. Mondellini, S. Pizzagalli, L. Greci, and M. Sacco. Assessment of an immersive virtual supermarket to train post-stroke patients: A pilot study on healthy people. In *Augmented Reality, Virtual Reality, and Computer Graphics: 6th International Conference, AVR 2019, Santa Maria al Bagno, Italy, June 24–27, 2019, Proceedings, Part I 6*, pages 313–329. Springer, 2019

12. S. Pizzagalli, D. Spoladore, S. Arlati, M. Sacco, and L. Greci. Hic: An interactive and ubiquitous home controller system for the smart home. In *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)*, pages 1–6. IEEE, 2018

Elulookirjeldus

1. Isikuandmed

Nimi	Simone Luca Pizzagalli
Sünniaeg ja -koht	01.02.1983, Merate, Italy
Kodakondsus	Itaalia

2. Kontaktandmed

Aadress	Tallinna Tehnikaülikool, Inseneriteaduskond, Mehaanika ja tööstustehnika Instituut, Ehitajate tee 5, 19086 Tallinn, Estonia
E-post	simone.pizzagalli@taltech.ee

3. Haridus

2020– . . .	Tallinna Tehnikaülikool, Inseneriteaduskond, Mehaanika ja tööstustehnika instituut, doktorant
2015–2016	Italian National Research Council, Politecnico di Milano, Design for All: Advanced Human Machine Interfaces
2005–2007	Delft University Of Technology, Faculty of Architecture, Architecture, urban planning and building sciences, MSc, *cum laude*
2008–2011	Politecnico di Milano, Faculty of Architecture, Architecture and urban planning, BSc

4. Keelteoskus

itaalia keel	emakeel
inglise keel	kõrgtase
hispaania keel	algtase

5. Teenistuskäik

2020– . . .	Tallinna Tehnikaülikool, Inseneriteaduskond, Mehaanika ja tööstustehnika instituut, doktorant-nooremteadur
2017–2020	Italian National Research Council, STIIMA, Lecco, Italy, nooremteadur
2018–2018	Soongsil University, Seoul, South Korea, Külalisprofessor
2016–2016	Politecnico di Milano, Milan, Itaalia, Teadustöö Praktika
2012–2014	Osar n.v., Antwerp, Belgia, Disainer, Teadur
2010–2014	Menabo' architecture, Rotterdam, Madalmaad, Disainer
2008–2010	De Werff Architectuur, The Hague, Madalmaad, Disainer
2007–2007	ZUS, Zones Urbaines Sensibles, Rotterdam, Madalmaad, Disainer Praktika

6. Vabatahtlik töö

2021–2023	NOOR MTÜ, Eesti.
2014–2016	Piccoli Artigiani. Osalusdisain ja planeerimine, sisekujunduse renoveerimine kasutatud materjalidega, Itaalia.
2010–2014	Asutaja POST—OFFICE arhitektuursest platvormist Rotterdamis: hüljatud asukoha renoveerimine ja aktiveerimine Rotterdami põhjaosas; osalusdisaini tegevused ja projektid; arhitektuurikultuuri ja käsitöö edendamine, Madalmaad.

7. Autasud

- 2023 Parim esineja, 2023 International Conference on Computers and Automation.
- 2023 Parima demo auhind, EUROXR 2023: ROS based XR interface for robot path-planning and teleoperation.
- 2018 H4O Hackathon for Ophthalmology Winner of category CARE / Winner of Fondazione Cottino Price Projektiettepanek seadmele, mille eesmärk on parandada nägemispuudega inimeste hoolduslahendusi.
- 2011 Parima lõputöö esimese koha auhind, Archiprix International 2011, New York.

8. Teadustöö põhisuunad

- Human Machine Interaction
- Extended Reality
- User Experience

9. Teadustegevus
Avaldatud tööd, mis ei sisaldu lõputöös.

1. W. Terkaj, E. Pessot, V. Kuts, Y. Bondarenko, S. L. Pizzagalli, and K. Kleine. A framework for the design and use of virtual labs in digital engineering education. In *AIP Conference Proceedings*, volume 2989-1. AIP Publishing, 2024
2. Y. Bondarenko, S. L. Pizzagalli, V. Kuts, E. Petlenkov, and T. Otto. Digital twin based learning from demonstration system for industrial robots. In *ASME International Mechanical Engineering Congress and Exposition*, volume 87608, page V003T03A081. American Society of Mechanical Engineers, 2023
3. Y. Bondarenko, V. Kuts, S. Pizzagalli, K. Nutonen, N. Murray, and E. O'Connell. Universal xr framework architecture based on open-source xr tools. In *International XR Conference*, pages 87–98. Springer, 2023
4. M. Mondellini, S. Arlati, H. Gapeyeva, K. Lees, I. Märitz, S. L. Pizzagalli, T. Otto, M. Sacco, and A. Teder-Braschinsky. User experience during an immersive virtual reality-based cognitive task: a comparison between estonian and italian older adults with mci. *Sensors*, 22(21):8249, 2022
5. K. Kalda, S.-L. Pizzagalli, R.-M. Soe, R. Sell, and M. Bellone. Language of driving for autonomous vehicles. *Applied Sciences*, 12(11):5406, 2022
6. D. Diachenko, A. Partyshev, S. L. Pizzagalli, Y. Bondarenko, T. Otto, and V. Kuts. Industrial collaborative robot digital twin integration and control using robot operating system. *Journal of Machine Engineering*, 22, 2022
7. V. Kuts, Y. Bondarenko, M. Gavriljuk, A. Paryshev, S. Jegorov, S. Pizzagall, and T. Otto. Digital twin: Universal user interface for online management of the manufacturing system. In *ASME International Mechanical Engineering Congress and Exposition*, volume 85567, page V02BT02A003. American Society of Mechanical Engineers, 2021

8. R. Nossa, M. Porro, O. Picciolini, M. Malosio, S. Pittaccio, M. Valoriani, V. Asperti, F. Clasadonte, L. Oliveto, M. Mondellini, et al. What children with neuromotor disabilities need to play with technological games. *Applied Sciences*, 11(21):9886, 2021

9. L. Ordile, Y. Bondarenko, S. L. Pizzagalli, V. Kuts, and T. Otto. Augmented reality interface for industrial robot control and teleoperation. *nil e H ja K*, 2021

10. M. Mondellini, S. Arlati, S. Pizzagalli, L. Greci, M. Sacco, G. Ferrigno, and M. Sacco. Assessment of the usability of an immersive virtual supermarket for the cognitive rehabilitation of elderly patients: a pilot study on young adults 2018 ieee 6th international conference on serious games and applications for health (segah), 1–8 https://doi. org/10. 1109. *SeGAH*, 84013:13, 2018

11. M. Mondellini, S. Pizzagalli, L. Greci, and M. Sacco. Assessment of an immersive virtual supermarket to train post-stroke patients: A pilot study on healthy people. In *Augmented Reality, Virtual Reality, and Computer Graphics: 6th International Conference, AVR 2019, Santa Maria al Bagno, Italy, June 24–27, 2019, Proceedings, Part I 6*, pages 313–329. Springer, 2019

12. S. Pizzagalli, D. Spoladore, S. Arlati, M. Sacco, and L. Greci. Hic: An interactive and ubiquitous home controller system for the smart home. In *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)*, pages 1–6. IEEE, 2018

ISSN 2585-6901 (PDF)
ISBN 978-9916-80-179-6 (PDF)