Universidade de Aveiro
2024

João Luís Fernandes Clemente

Uso de um braço robótico para auxiliar cenários de Colaboração Remota apoiada por Realidade Mista

Using a Robotic Arm to Assist during Scenarios of Remote Collaboration supported by Mixed Reality

Universidade de Aveiro
2024

João Luís Fernandes Clemente

Uso de um braço robótico para auxiliar cenários de Colaboração Remota apoiada por Realidade Mista

Using a Robotic Arm to Assist during Scenarios of Remote Collaboration supported by Mixed Reality

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Robótica e Sistemas Inteligentes , realizada sob a orientação científica do Doutor Bernardo Marques, Professor auxiliar do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro, e do Doutor Eurico Pedrosa, Professor auxiliar do Departamento de Eletrónica, Telecomunicacoes e Informática da Universidade de Aveiro.

ponho aqui algo?

o júri / the jury

presidente / president	Prof. Doutor João Antunes da Silva professor associado da Universidade de Aveiro
vogais / examiners committee	Prof. Doutor João Antunes da Silva professor associado da Universidade de Aveiro
	Prof. Doutor João Antunes da Silva professor associado da Universidade de Aveiro
	Prof. Doutor João Antunes da Silva professor associado da Universidade de Aveiro
	Prof. Doutor João Antunes da Silva professor associado da Universidade de Aveiro
	Prof. Doutor João Antunes da Silva professor associado da Universidade de Aveiro

agradecimentos / acknowledgements

Esta dissertação consiste no culminar de um percurso académico que se estendeu ao longo de 17 anos.
Deste modo, gostaria de começar por agradecer à minha família, especialmente à minha mãe e ao meu pai por me terem proporcionado esta aventura, bem como, por terem celebrado todas as minhas conquistas.
Em segundo lugar, gostaria de agradecer aos meus orientadores, Professor Doutor Bernardo Marques e Professor Doutor Eurico Pedrosa, pelo apoio e orientação na conceção deste trabalho, bem como pelas diversas horas nas quais se dispuseram a ajudar-me, de modo a poder levar este projeto a bom porto. Gostaria também de agradecer ao Doutor João Alves, por me ter transmitido o seu conhecimento e experiência, fundamentais para a realização deste trabalho.
Por fim, mas não menos importante, gostaria de agradecer a todos os amigos que fiz ao longo deste percurso, tendo um especial carinho pela "familia" que criei em Aveiro, por terem partilhado comigo belos momentos e pela constante presença e apoio nos últimos 5 anos.
A todos aqueles que se cruzaram comigo e de algum modo contribuíram para o meu crescimento pessoal e académico, o meu muito obrigado.

palavras-chave

texto livro, arquitetura, história, construção, materiais de construção, saber tradicional.

resumo

Um resumo é um pequeno apanhado de um trabalho mais longo (como uma tese, dissertação ou trabalho de pesquisa). O resumo relata de forma concisa os objetivos e resultados da sua pesquisa, para que os leitores saibam exatamente o que se aborda no seu documento.
Embora a estrutura possa variar um pouco dependendo da sua área de estudo, o seu resumo deve descrever o propósito do seu trabalho, os métodos que você usou e as conclusões a que chegou.
Uma maneira comum de estruturar um resumo é usar a estrutura IMRaD. Isso significa:

- Introdução
- Métodos
- Resultados
- Discussão

Veja mais pormenores aqui:
`https://www.scribbr.com/dissertation/abstract/`

keywords textbook, architecture, history, construction, construction materials, traditional knowledge.

abstract An abstract is a short summary of a longer work (such as a thesis, dissertation or research paper).
The abstract concisely reports the aims and outcomes of your research, so that readers know exactly what your paper is about.
Although the structure may vary slightly depending on your discipline, your abstract should describe the purpose of your work, the methods you've used, and the conclusions you've drawn.
One common way to structure your abstract is to use the IMRaD structure. This stands for:

- Introduction
- Methods
- Results
- Discussion

Check for more details here:
`https://www.scribbr.com/dissertation/abstract/`

reconhecimento do uso de ferramentas de AI

Reconhecimento do uso de tecnologias e ferramentas de Inteligência Artificial (IA) generativa, softwares e outras ferramentas de apoio.
Reconheço a utilização do ChatGPT 4 (Open AI, https://chat.openai.com) para melhorar a escrita académica e para fornecer sugestões de código e assistência no desenvolvimento do software.

I acknowledge the use of ChatGPT 4 (Open AI, https://chat.openai.com) for refining academic language and offering code suggestions, aiding in the development of the software components.

Contents

List of Figures

List of Tables

Acronyms

AI	Artificial Intelligence
AR	Augmented Reality
AR	Augmented Reality
AV	Augmented Virtuality
CPS	Cyber-Physical System
CPSS	Cyber-Physical Systems
DR	Digital Reality
DT	Digital Twin
FOV	Field of View
HHD	Handheld Device
HHDs	Handheld Devices
HMDs	Head-Mounted Displays
HMI	Human-Machine Interaction
HRC	Human-Robot Collaboration
HRCp	Human-Robot Cooperation
HRCx	Human-Robot Coexistence
HRI	Human-Robot Interaction
IDE	Integrated Development Environment
IoT	Internet of Things
IP	Internet Protocol
ML	Machine Learning
MR	Mixed Reality
ROS	Robot Operating System
SDK	Software Development Kit
SMEs	Small and Medium Enterprises
TCP	Transmission Control Protocol
UI	User Interface
URDF	Unified Robot Description Format
VR	Virtual Reality

CHAPTER 1

Introduction

This chapter establishes the groundwork for this dissertation by outlining the motivation, goals, and structure of the research. It discusses the importance of enhancing remote collaboration through Mixed-Reality and Digital Twin technologies. It then provides an overview of the dissertation's structure, guiding the reader through each stage, from theoretical foundations and technical implementation to in-depth discussion and proposed directions for future work.

1.1 Motivation

The First Industrial Revolution, powered by steam engines, paved the way for subsequent revolutions driven by electricity, automation, machinery, and the internet. Each revolution introduced groundbreaking technologies that reshaped industries, emphasizing the companies' need to prioritize reskilling and upskilling their workforce.

In recent years, Industry 4.0 has marked a paradigm shift toward the digitization of manufacturing processes. By integrating technologies such as the Internet of Things (IoT), Artificial Intelligence (AI), and automation, Industry 4.0 has fostered highly interconnected and intelligent manufacturing ecosystems. These advancements have led to a notable decrease in the reliance on human labor, as automation takes center stage in many operations. However, as factories become increasingly autonomous, the unique cognitive and adaptive capabilities that humans bring to complex decision-making, creativity, and problem-solving remain irreplaceable, and experts argue that complete removal of humans from the manufacturing processes is not feasible. Instead, there is a growing emphasis on fostering collaborative partnerships between humans and intelligent machinery [1].

As collaborative environments evolve, robots have become indispensable in various domains, leading to increased complexity in these scenarios. Therefore, advanced solutions are needed to enhance HRC. For instance, in flexible manufacturing systems, robots must adapt to frequent changes in production tasks, requiring advanced cognitive capabilities to handle non-repetitive operations. Additionally, in collaborative assembly lines, robots need to interact safely and efficiently with human workers, often in unstructured environments where unpredictable human movements pose significant safety and operational challenges. One promising approach consists on integrating Mixed Reality (MR) technologies as a medium for collaboration, encompassing Virtual Reality (VR) for the remote user and Augmented Reality (AR) for the on-site one, by blending the physical and digital world. By providing immersive experiences that transcend traditional reality and overcome geographical constraints, this relationship enables real-time collaboration among individuals from different locations [2]. * TODO: add more references - por 3/4 aqui - verificar melhor ao longo do resto do texto, 4 que façam sentido

However, the potential of MR to enhance remote collaboration is currently hindered by several critical limitations. These include, not only limited perspective and context capture, which impede remote collaborators' understanding and decision-making capabilities, as well as a lack of multisensory data collection, restricting comprehensive environmental comprehension. Additionally, MR interaction with physical objects often lacks the precision required for detailed tasks, particularly in dynamic scenarios. These challenges diminish the effectiveness of MR in facilitating thorough context sharing and impact the overall efficiency and safety of collaborative tasks. * TODO: add more references - verificar melhor ao longo do resto do texto algumas que façam sentido aqui

1.2 Goals

The primary goal of this dissertation is to enhance remote collaboration between human operators by utilizing a robotic arm (UR10e) and MR technologies. This framework enables dynamic and immersive collaboration, where both on-site and remote participants can interact with both the robot and the shared environment in real time.

According to the collaborative element being addressed, namely:

- **On-Site Interaction:**
  - Enable dynamic and real-time robot manipulation through AR within the designated environment.
  - Infer robot state
  - Visualize and interact with the robot's workspace through AR safety-zones, enhancing the user's situational awareness.
  - Utilize Handheld Devices (HHDs), such as tablets or smartphones, to share live views of the surroundings, allowing remote collaborators to gain a comprehensive understanding of the collaborative space.
- **Remote Visualization and Interaction:**
  - Provide remote member with a foundational UI interface, such as a laptop screen, to visualize the collaboration scenario and task context.
  - Establish a bidirectional communication, enabling remote operation of the robot arm via the MR application, enhancing the user's ability to interact with the on-site physical space of his/her counterpart.

1.3 Thesis Structure

This dissertation is structured across six chapters, each representing a logical progression in the development of this work within the context of HRC and Industry 5.0.

The first chapter introduces the project by framing the motivation behind this work as well as defining the baseline for this dissertation's development. Then, a thorough state-of-the-art review is presented, focusing on key concepts such as Industry 5.0, HRC, Collaborative Robots, Digital Reality (DR), and DT. It also includes a description of case studies involving MR-DT solutions in industrial settings. The third chapter discusses several implementation tools utilized during the project development, such as Unity, Vuforia, and Robot Operating System (ROS). Chapter four presents the established MR framework foundational to this dissertation, elaborating on specific implemented features and distinguishing those tailored for remote users from those intended for on-site collaborators. Chapter five discusses the developed functionalities and the challenges encountered during implementation. Despite not having conducted user studies, potential industrial application scenarios are explored, illustrating possible real-world implementations. Finally, chapter six concludes the dissertation with a summary of key achievements and proposed directions for future development.

CHAPTER 2

State of Art

This chapter reviews key concepts driving the transition from Industry 4.0 to Industry 5.0, focusing on how Human-Robot Collaboration is evolving in industrial settings. It begins with a comparison between the automation-driven approach of Industry 4.0 and the more human-centered framework of Industry 5.0. It further explores technologies like collaborative robots, Digital Twins, and Digital Realities, discussing their potential in creating safe and effective collaborative environments. The chapter also highlights challenges, research advancements, and future directions in these fields.

2.1 Key Drivers of Industry 5.0

With the advent of Industry 4.0 and the emerging concept of Industry 5.0, the industry environment has witnessed significant transformations. Understanding this progress is crucial to contextualize the technological advancements and the shift towards more human-centric manufacturing processes.

2.1.1 Industry 4.0: The Fourth Industrial Revolution

Industry 4.0 represents the integration of cutting-edge digital technologies into manufacturing processes, leading to the emergence of smart factories. It leverages advanced systems such as Cyber-Physical Systems (CPSS), the IoT, robotics, AR, simulation, cloud computing, and big data analytics, as illustrated in Figure 2.1. This paradigm signifies a fundamental shift towards interconnected, intelligent, and digitally-driven manufacturing ecosystems, revolutionizing the way products are conceived, manufactured, and delivered, while enhancing production efficiency, flexibility, and innovation [3, 4].

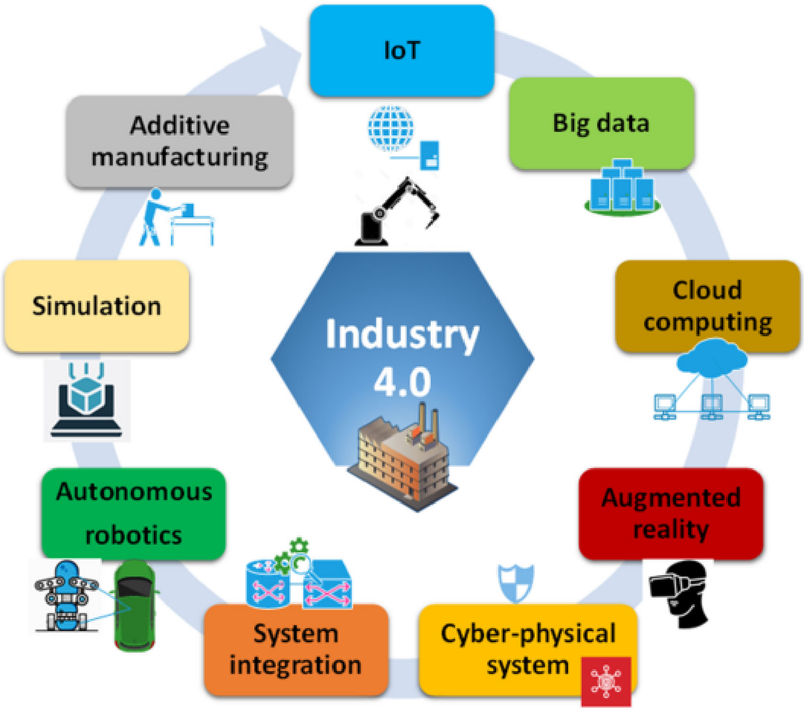


Figure 2.1: Key enabling technologies in Industry 4.0 [4]

With advancements in AI, industrial processes can achieve unprecedented performance levels, often exceeding human capabilities. These AI-driven systems enable robots to perform tasks that may be too hazardous, complex, or delicate for humans, such as handling dangerous materials or managing microscopic elements. Despite this extraordinary potential, it is important to recognize that current industrial robots are not as "smart" as humans in many contexts and, even though these robots are capable of performing highly skilled tasks, they frequently operate under strict, pre-programmed limits [4].

Although Industry 4.0 has undoubtedly increased productivity, flexibility, and automation in industrial environments, it has also led to concerns regarding the diminishing role of human operators. This relentless push towards full automation has, in some cases, reduced human involvement in critical decision-making processes, leading to a more machine-centric production landscape [5].

2.1.2 Industry 5.0: Reintegrating the Human Element

Approximately a decade after the launch of Industry 4.0, the European Commission introduced the Industry 5.0 concept in response to new societal challenges [6]. The growing concerns about the exclusion of human operators in Industry 4.0 systems, coupled with the limitations of full automation, paved the way for this new industrial paradigm. Industry 5.0 seeks to reintroduce the human element into industrial ecosystems, emphasizing greater human involvement in manufacturing processes [7].

The main goal consists on combining the strengths of humans and machines to achieve more sustainable, efficient, and human-centered production systems. This shift reflects the realization that, while machines excel at repetitive, dangerous, or complex tasks, humans provide irreplaceable creativity, adaptability, and problem-solving abilities [8]. Industry 5.0 aims to strike a balance between technological advancement and human-centric values, fostering environments where humans work alongside advanced technologies to achieve greater societal and environmental outcomes [5].

Recognizing that humans and machines each possess distinct strengths that can complement one another, the following key technological drivers of Industry 5.0 build upon the advancements of Industry 4.0 [8]:

- **Collaborative Robots (Cobots):** are engineered to ensure safe, collaborative operation alongside human workers, facilitating not only intuitive interactions but also fostering efforts that leverage the unique strengths of both humans and robots. Their integration is driven by the need to create systems that enable seamless, user-friendly HRC, in full alignment with the guiding principles of Industry 5.0. This paradigm shift redefines traditional employment roles by emphasizing HRI, with a focus on communication and coordination with robotic systems and advanced AI.
- **Digital Twins:** represent a pivotal technological advancement in Industry 5.0. They provide visual models that enhance comprehension and facilitate the evaluation of goods, processes, and production systems. By allowing real-time monitoring and simulation, DT help optimize manufacturing processes, bridging the gap between the virtual and physical worlds.
- **Human-Centric Automation:** Emphasis is placed on using technology to augment human capabilities rather than replace them, fostering a more inclusive, creative, and flexible manufacturing environment. This approach ensures that technology empowers human workers, enabling them to focus on tasks requiring intuition and creativity.
- **Advanced Human-Machine Interfaces:** The development of intuitive interfaces, by integrating technologies such as AR and VR, facilitates better communication between humans and machines. These interfaces allow for more natural interactions, improving understanding and efficiency in collaborative tasks.
- **Artificial Intelligence and Cognitive Computing:** These technologies continue to evolve, enabling robots and automation systems to work alongside humans in ways that enhance productivity without fully replacing them. AI allows for more intuitive Human-Machine Interaction (HMI), where machines can understand and respond to human needs more effectively.
- **Sustainable and Resilient Manufacturing:** Industry 5.0 also focuses on sustainability and resilience, integrating environmental considerations into manufacturing processes. This includes optimizing resource usage and reducing waste, aligning technological advancement with ecological responsibility.

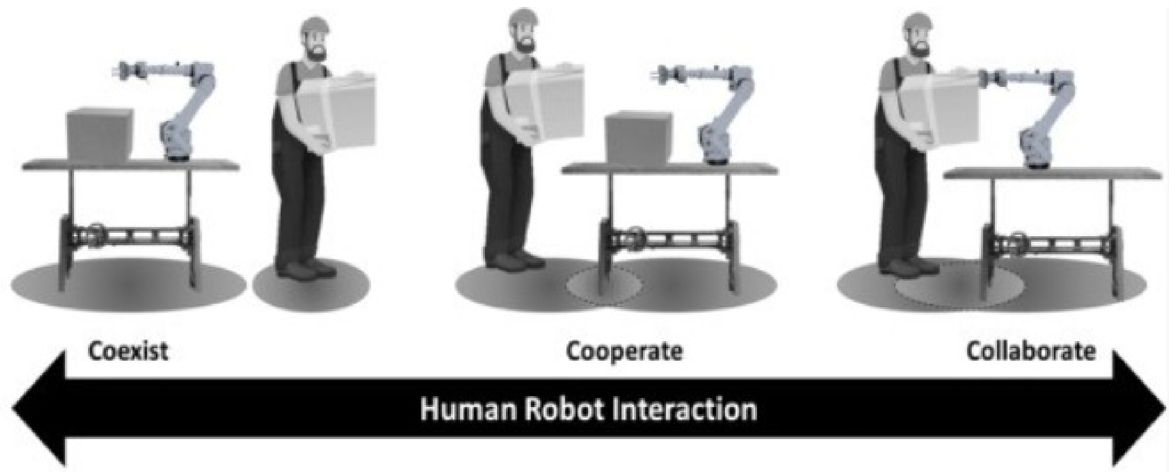
By integrating these key drivers, Industry 5.0 addresses the challenges identified in Industry 4.0, promoting harmonious collaboration between humans and machines. This synergy aims to enhance productivity while preserving the unique contributions of human workers, ultimately leading to more innovative, sustainable, and human-centered industrial practices.

2.2 Human-Robot Collaboration

The field of HRI is dedicated to examining the interactions and coexistence of humans and robots in shared spaces, whose objective consists on enhancing these interactions by designing robots that are safe, effective and compatible for assisting and cooperating with humans in diverse roles, rather than replacing them [9]. This involves developing robots that are, not only autonomous, but also capable of understanding and communicating with humans, as well as predicting human-behavior and learning from human feedback.

However, HRI can be broken down into different forms of interaction, whose categorization is based on various factors that define how humans and robots share the workspace. This disctinction is represented in the Figure 2.2 and can be broken down into **Coexistence**, **Cooperation** and **Collaboration** [10]. Each one can be distinguished by the degree of interaction and task sharing between humans and robots:

* TODO: i tried to edit it and center it, but it lost resolution. i will try to make it better



Space	Separated	Shared	Shared
Tasks	Unrelated	Linked	Shared
Contact	Impossible	Possible	Desired
Context Awareness	Not necessary	Minimal	Necessary

Figure 2.2: Different levels of HRI [10]

- **Human-Robot Coexistence (HRCx)**: In this form of interaction, humans and robots operate in the same environment but perform entirely independent tasks without interaction. The workspace is separated, there is no contact, and context awareness is unnecessary since the tasks are unrelated. Usually this interaction does not involve synchronized work or communication between the two parties.
- **Human-Robot Cooperation (HRCp)**: Here, humans and robots share the workspace and work on linked tasks, towards a common goal. Advanced technologies such as sensors or machine vision may be used to detect and prevent collisions. Contact is possible, though not essential and actions are largely independent with occasional coordinated efforts.

- **Human-Robot Collaboration (HRC)**: represents the most advanced and integrated form of HRI. In HRC, humans and robots not only share a common workspace but also actively collaborate on shared objectives. This collaboration can involve direct physical contact, such as the joint manipulation of objects, or non-physical interaction, including verbal communication, gestures, or pattern recognition. Within such environments, humans often handle tasks that require fine motor skills, decision-making, or creative problem-solving, while robots take on repetitive, strenuous, or hazardous activities, ensuring efficiency and safety. This synergy enhances productivity by leveraging the unique strengths of both humans and robots, creating a dynamic partnership where each complements the other's capabilities.

In the bottom part of the Figure 2.2 there is a table that further breaks down these distinctions, highlighting key factors like space, task relationship, possibility of contact, and the need for context awareness. This gradient from coexistence to collaboration shows the increasing complexity and interdependence in HRI, as technology evolves to make robots more capable partners in industrial and service environments.

Below, the Figure 2.3 illustrates a HRC workspace, showcasing the interaction between a human worker and a robot within a shared environment. The workspace is divided into distinct yet overlapping areas: the robot's work envelope and the human's work envelope. These areas reflect the respective tasks of each party, with the robot likely performing repetitive or automated tasks, such as material handling within the feeding system, while the human focuses on more intricate tasks requiring dexterity and decision-making.

The overlapping shared workspace demonstrates the core principle of HRC, where humans and robots work together toward common goals, necessitating real-time coordination and communication. In this setup, advanced sensing technologies or machine vision are essential to ensure safety and prevent collisions, allowing both the human and robot to operate efficiently within close proximity.

This image underscores how robots, rather than replacing humans, complement human skills by taking on routine, physically demanding tasks, while humans contribute with their cognitive abilities. This partnership reflects the broader vision of Industry 5.0, where human creativity and robotic precision are combined to create adaptable, human-centered industrial processes, enhancing both efficiency and safety in collaborative environments.

These new robots featuring intelligent sensing and vision systems, envisioned to integrate the production line, are called "cobots". They represent the alternative to full automation, since industry specialists have stated it is not possible to completely remove the human within the manufacturing environment [1].

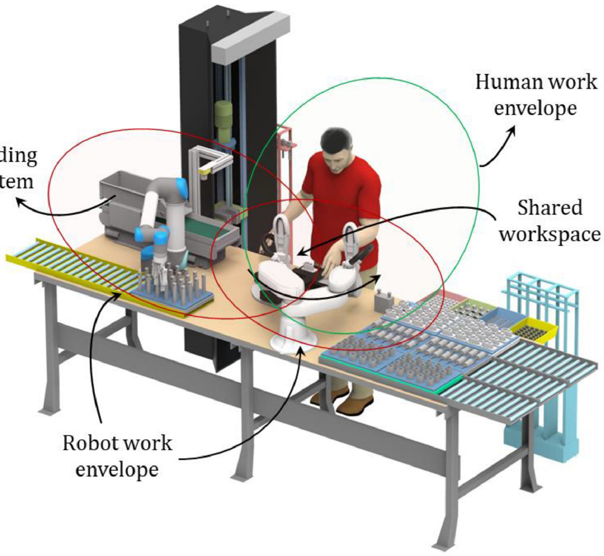


Figure 2.3: Workstation example enabling collaboration between human and robot while sharing the same physical space [11]

2.3 Collaborative Robots (Cobots)

The concept of collaborative robots, or "cobots," was first introduced by J. Edward Colgate and Michael Pashkin in 1996 [12], laying the foundation for practical applications in HRC. Cobots are designed to physically interact with humans in shared workspaces, without the need for protective barriers typical in traditional robotic systems. This innovation paved the way for a new category of robots that excel in adaptability and flexibility, although effective use requires a deep understanding of their unique characteristics [13].

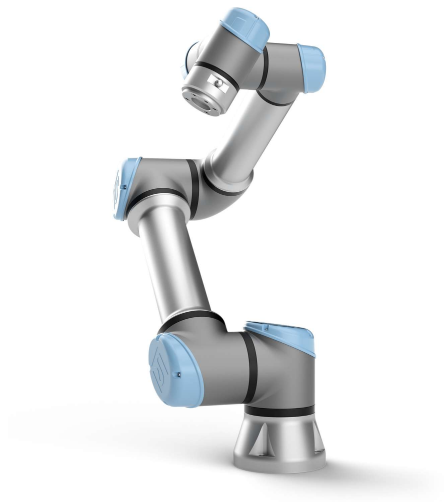


Figure 2.4: Universal Robots UR5 cobot

Commercializing cobots began in 2008 with the release of the UR5 model by Universal Robots, marking a breakthrough in HRC technologies and facilitating the integration of cobots into industrial workflows [13]. The UR5, shown in Figure 2.4, represents a pivotal development in collaborative robotics, as it enabled quicker and more cost-effective adaptation of industrial layouts.

Unlike traditional industrial robotic systems, which require extensive safety guarding and consequently reduce flexibility while increasing both costs and spatial demands, cobots present a solution tailored to the current market's demand for shorter lead times and mass customization, particularly for Small and Medium Enterprises (SMEs) [14].

This cobots' emergence represents a paradigm shift in industrial automation, emphasizing HRC over the traditional model of robotic isolation. These facilitate direct physical interaction between humans and machines while being designed for intuitive use, enabling even non-experts to reprogram them effortlessly [15]. By leveraging the complementary strengths of human cognitive capabilities and robotic precision, cobots offer substantial productivity gains and reduced operational costs.

Cobots differentiate themselves from traditional industrial robots by prioritizing safety, ergonomics, and user accessibility. Unlike conventional robots that require extensive safety enclosures, cobots are equipped with advanced features such as force and torque sensors, vision systems, and anti-collision mechanisms. These capabilities enable them to operate safely in close proximity to humans without the need for restrictive barriers [16]. The inherent design of cobots supports flexibility and ease of deployment, avoiding the high costs and complexity associated with retrofitting traditional robotic systems for similar functionality.

The adoption of cobots in industrial settings is driven by a combination of economic, operational, and health-related factors:

- **Cost Efficiency:** Cobots can significantly reduce labor costs by performing repetitive tasks, thereby lowering direct unit production costs compared to traditional automation solutions [17].
- **Enhanced Workplace Safety:** Their design minimizes occupational hazards, which leads to improved worker safety and health, addressing ergonomic challenges in manual labor.
- **Spatial Efficiency:** The compact and flexible nature of cobots allows them to be easily relocated and reconfigured within different production areas, optimizing factory space utilization [18].

These attributes are particularly beneficial in high-risk applications and industries that demand frequent changes in production layouts, such as electronics, automotive, and aerospace manufacturing.

When assessing the applicability of cobots versus traditional robots, several distinctions emerge. Cobots excel in tasks that require adaptability and human-like dexterity, such as assembly, placement, handling, and quality inspection. Their versatility and ease of integration make them suitable for low-volume, high-mix production environments, where agility is crucial. However, traditional robots retain an advantage in scenarios that demand high payload capacity, speed, and precision, particularly in heavy-duty manufacturing applications [13].

Despite their advantages, integrating cobots still encounters challenges for some use-cases. Issues such as interoperability with legacy systems, programming complexity for non-standard tasks, and optimizing cobot performance in dynamic environments require further research and development. Additionally, advancements in AI and Machine Learning (ML) could unlock new capabilities for cobots, enabling them to autonomously adapt to changing tasks and work conditions, thereby extending their utility beyond predefined, structured environments.

Moreover, MR technologies are becoming increasingly pivotal in advancing cobot integration, offering immersive, real-time interfaces that enhance HRC. By overlaying digital information onto the physical workspace, MR facilitates better situational awareness and task execution for both operators and cobots. Coupled with DT systems, MR enables real-time monitoring and control of cobots in dynamic and remote settings, enhancing operational flexibility. Integrating MR into cobot applications thus allows for more intuitive interactions, improved safety, and higher precision in collaborative environments. As part of Industry 5.0, these technologies collectively contribute to creating more responsive, human-centered manufacturing systems.

2.4 Digital Realities

DRs! (**DRs!**) encompass a wide spectrum of technologies that merge virtual elements with real-world environments to varying extents. In 1994, Milgram and Kishino introduced the Reality-Virtuality Continuum, a theoretical framework that characterizes the progression from a purely physical environment to a fully virtual one, as illustrated in Figure 2.7 [19]. This continuum is divided into four principal stages: Reality, AR, Augmented Virtuality (AV), and VR.

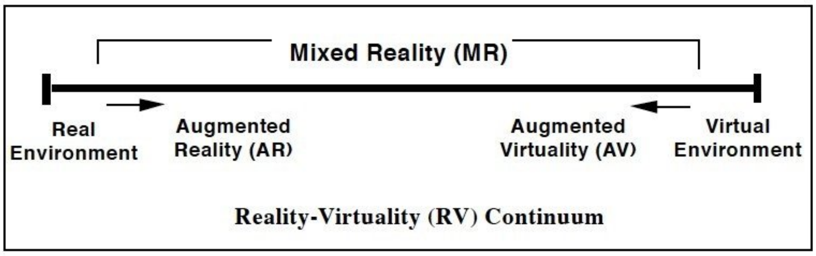


Figure 2.5: Reality-Virtuality Continuum [19]

In this continuum, Reality represents the perception of an unaltered physical environment, devoid of any virtual modifications. As we progress along the continuum towards the virtual side, different digital realities offer increasingly immersive experiences by blending or replacing real-world content with virtual elements.

2.4.1 Augmented Reality

AR enhances a user's interaction with their physical environment by overlaying dynamic digital content, such as 3D objects, information layers, or media, onto the real world [20]. The main goal of AR is to seamlessly integrate virtual objects with the user's surrounding physical context, facilitating real-time interaction between the virtual and physical realms allowing users to experience both virtual and real entities as coexisting within the same space, generating a cohesive and interactive environment [21].

Figure 2.6 [1] provides a clear representation of AR being employed in an industrial maintenance scenario. The user views real-time digital overlays on a physical machine, providing essential information such as the machine's current status and highlighting critical components for maintenance or repair. This visual guidance significantly enhances the user's understanding of tasks by seamlessly blending virtual instructions with the real environment, thereby improving efficiency and reducing the potential for human error.

Figure 2.6: Augmented Reality application of industry overlayed information in order to assist the operator

Achieving this level of integration requires accurate spatial registration, a process that ensures that virtual elements are properly aligned with real-world objects in both location and scale. The spatial coherence between the two realities is critical for creating an effective AR experience, where virtual objects respond to changes in the environment and user interaction in real-time.

Various AR devices are employed to deliver these experiences, including AR-Head-Mounted Displays (HMDs), tablets, HHDs, projectors, and see-through VR headsets with built-in cameras. Each device offers different degrees of environmental awareness and interaction capabilities, such as hand tracking and holographic projection.

* TODO: add more references to the above 2 paragraphs

2.4.2 Virtual Reality

Within the Reality-Virtuality Continuum proposed by Milgram and Kishino, VR occupies the extreme end of the spectrum, representing a complete substitution of a user's perception of the physical world with a fully immersive synthetic environment. In this state, the user is entirely isolated from their real surroundings, perceiving only the artificially constructed virtual environment, typically presented through a range of immersive devices such as HMDs [19]. Figure 2.7 [2] illustrates this immersive experience where the user interacts with a virtual

[1] `https://www.cad-schroer.de/news-events/artikel/mixed-reality-augmented-reality-virtual-reality-definitio` Acessed: 2024-10-19

[2] `https://en.wikipedia.org/wiki/Immersion_%28virtual_reality%29#/media/File:Reality_check_ESA384313.jpg`

environment through VR equipment.

Figure 2.7: User immersed in a virtual environment using a VR headset, demonstrating complete isolation from the physical world.

Modern VR systems achieve this full immersion by leveraging advanced HMDs, which present stereoscopic images directly to the user's eyes through built-in displays or projection systems. These devices often incorporate additional features like head tracking, enabling the user's head movements to influence their viewpoint in the virtual environment, further enhancing the sense of immersion and presence [2]. Some systems also include positional tracking through external sensors or inside-out tracking via integrated cameras, allowing users to physically navigate virtual spaces, augmenting both interaction fidelity and spatial awareness.

VR is particularly effective in applications that require the user to be completely enveloped in an artificial environment, thus enabling the simulation of real-world scenarios, historical reconstructions, or entirely imaginative worlds. This sense of "presence," wherein users perceive the virtual environment as real, is fundamental to VR's efficacy across various domains, including gaming, education, training, and simulation. Additionally, VR's potential to create deeply immersive and isolated experiences makes it especially valuable in fields like remote collaboration, where users can interact with simulated environments or models that are otherwise inaccessible [22].

* TODO: add more references here

2.4.3 Mixed Reality

MR continues to elude a universally accepted definition, with interpretations diverging significantly across academic and industrial domains. According to the Milgram and Kishino Reality-Virtuality Continuum, depicted in Figure 2.7, MR occupies a transitional space between AR and AV, bridging the two concepts [19]. Expanding on this, Microsoft's MR spectrum defines MR as spanning a range of technologies, from AR (where physical reality predominates, augmented with digital overlays) to AV (where the virtual environment dominates, supplemented by real-world data) [23].

* TODO: ADD REFERENCES IN BELOW TEXT In MR environments, digital and physical elements coexist and interact in real time, creating a dynamic interface where virtual and physical worlds seamlessly blend. This enables immersive, bi-directional interaction, where users engage with both digital objects and real-world elements, facilitating fluid communication between virtual entities and the physical environment. This integration enhances the user experience by enabling virtual objects to influence, and be influenced by, real-world contexts in a highly interactive manner, enabling new forms of collaboration, visualization, and interaction.

The inherent complexity of MR arises from the challenge of ensuring a natural and intuitive integration of digital and physical elements. This requires advanced environmental sensing, real-time data fusion, and contextual understanding to deliver interactions that appear natural to the user.

Despite considerable advancements, the definition of MR remains contested. Speicher et al. (2019) identified six competing notions of MR across both scholarly literature and industry practice, underscoring the fragmentation in its interpretation. Some experts argue that MR represents an enhanced form of AR, where users are not merely passive observers but active participants interacting with a responsive augmented space. In this interpretation, MR is seen as a "stronger" form of AR, exemplified by technologies such as Microsoft's HoloLens, where users can manipulate virtual elements within their physical environment [2].

Other perspectives view MR as a convergence of AR and VR, where the boundary between the real and virtual worlds is fluid and adaptable, creating immersive, hybrid experiences. For example, the widely known game Pokémon Go is sometimes cited as an MR application, where a VR-based digital environment enables users to interact with augmented digital elements such as Pokémons, overlaid onto the physical world [2].

However, the definition most relevant to this project's development emphasizes MR as a powerful medium for collaboration, enabling users to interact across different realities—whether physical, augmented, or virtual. In this context, MR facilitates shared experiences between users located in distinct environments. For example, a physical space visualized by an on-site AR user can be simultaneously recreated and experienced by a remote VR user, allowing real-time collaborative interactions between participants situated in different realities. This capacity for cross-reality collaboration enhances both the user's understanding of the workspace and the efficiency of the collaboration process.

An example of this collaborative interaction is depicted in Figure **??**, where a trainer (remote) and trainee (on-site) engage in a task that bridges physical and virtual worlds. In this scenario, the trainee uses VR technology to operate within a virtual environment, while the trainer remotely observes and provides guidance. Through this setup, the trainer can not only visualize the trainee's actions but also trigger events or scenarios via a dedicated interface, enhancing the learning experience. This type of interaction exemplifies the capabilities of remote training, where an expert can guide an on-site operator using real-time augmented indications displayed on AR glasses or tablets. By providing direct visual and auditory feedback, the remote trainer assists in troubleshooting, guiding the trainee through complex tasks or maintenance operations. This form of collaboration allows geographically distant participants to work synchronously in a highly interactive environment, increasing both efficiency and accuracy in training processes [24].

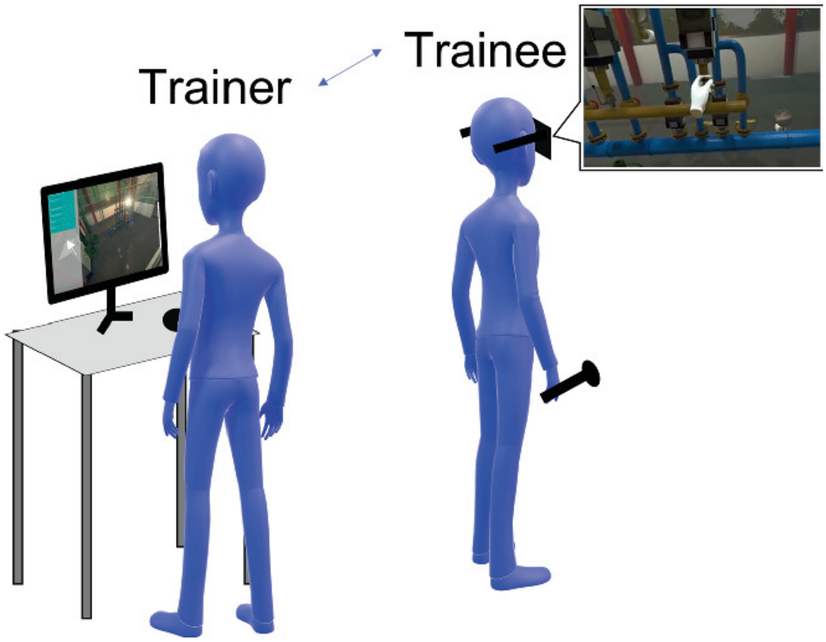


Figure 2.8: Collaborative setup for training in industrial scenarios using MR [24]

*TODO: Verify this part a bit better - clarify the distinction between the referenced collaboration scenario and the collaboration of this dissertation, can I say that the remote user sees is immersed into an VR environment that the on-site also sees? because the on-site member sees it via AR and not only VR?

Although this example highlights the collaborative potential of MR, it slightly differs from the ideal collaboration model developed in this project. Here, the focus is on enabling an **on-site** user to interact with the real world through AR—visualizing real-time data overlaid onto their physical environment—while the **remote** user experiences the same VR environment. This configuration allows the remote user to be fully immersed in a virtual reconstruction of the on-site environment, with the ability to monitor and collaborate with the on-site user in real-time. The goal is not only to enhance communication but to enable mutual interaction with a shared digital-physical environment, fostering more precise decision-making, particularly in complex, dynamic tasks. This setup, more reflective of the full potential of MR, aligns closely with Industry 5.0's vision of combining human expertise and advanced technology for seamless collaboration between humans and machines across different realities.

2.5 Digital Twins

Another relevant concept are **DTs!** (**DTs!**), which consist on sophisticated digital replicas of physical entities, allowing for the simulation, analysis, and control of systems within a digital framework. These digital counterparts have emerged as pivotal technologies in a variety of domains, particularly in enhancing HRC, as they offer real-time, interactive environments that mirror physical systems. The ability to replicate physical entities with high fidelity enables improved decision-making, operational efficiency, and flexibility across a wide range of industrial applications. *TODO: add a reference above

This DT concept was first introduced by NASA in the 1980s as part of its spacecraft monitoring systems, where virtual models were employed to replicate the conditions and behavior of spacecraft during missions. Over the past decades, advances in IoT, sensor technology, and computational power have significantly improved the capabilities of **DTs!**, moving beyond their original use case. Modern systems leverage real-time sensor data and advanced simulation techniques to enhance the accuracy and reliability of digital models, enabling more sophisticated predictions, real-time analytics, and simulations of complex systems [20].

In current manufacturing and industry scenarios, **DTs!** play a transformative role, particularly in smart manufacturing systems, where they enable detailed examination and prediction of the behavior of physical systems. This capability allows companies to optimize operations, reduce downtime, and improve overall system efficiency. Furthermore, in HRC, **DTs!** facilitate safer and more productive work environments by dynamically adjusting robotic movements and operations to better align with human needs, thereby enhancing ergonomic interactions and mitigating safety risks [25].

A prominent real-world implementation of this technology is seen in Singapore's Smart Nation Initiative, where the Land Transport Authority employs a DT to simulate and evaluate potential policy decisions before their implementation. This application exemplifies the wide-ranging potential of **DTs!** to support decision-making processes in urban planning, infrastructure management, and beyond [26]. As these technologies evolve, their applications in both academic research and industrial practice are rapidly expanding.

The academic landscape surrounding **DTs!** has seen extensive research exploring their versatility and potential, having been applied to a broad spectrum of areas, illustrating the profound impact of DTs on enhancing system efficiency, predictive maintenance, and overall operational performance [26–30]. *TODO: usar uma destas referencias para pôr no paragrafo introdutorio sobre os DTs.

Despite the growing prominence of **DTs!**, there is still no universally accepted formal definition of the concept. However, most scholars and industry experts concur that a DT is a type of Cyber-Physical System (CPS) consisting of three fundamental components: a physical system, a virtual model, and bidirectional communication between these two models [25, 31, 32]. This interaction is fundamental to the operation of true **DTs!**, allowing for a continuous feedback loop where changes in the physical world can inform the virtual model, and, in turn, decisions or optimizations made in the digital realm can directly influence the physical system.

In contrast to true **DTs!**, some critics argue that many commercially available implementations, such as those provided by companies like Siemens, represent "digital shadows" rather than full **DTs!**. The distinction lies in the capability for bidirectional communication. In many digital shadow systems, changes in the physical system are reflected in the virtual model, but there is no capacity for the virtual model to directly control or alter the physical system. This one-way communication limits the interactive and predictive capabilities that define a true DT [33].

To illustrate the distinction, Figure 2.9 presents a reference model of a DT, showcasing the bidirectional flow of information between physical and digital entities. This structure is essential for enabling real-time interaction and feedback, core characteristics that differentiate **DTs!** from digital shadows. True **DTs!** must facilitate a continuous, reciprocal exchange of data between physical and virtual domains, allowing the virtual model to reflect and affect the physical system [34].

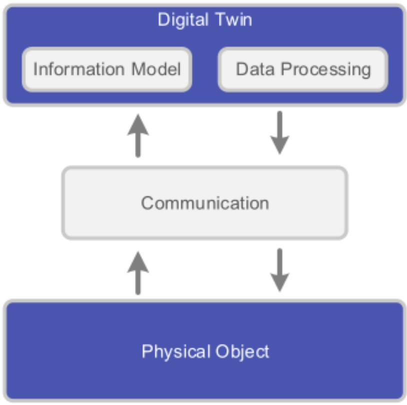


Figure 2.9: A Digital Twin reference model, emphasizing the importance of bidirectional communication between the physical object and its digital counterpart [34]

Bidirectional communication's importance in **DTs!** is further emphasized by Liu et al. [35], who argue that "a true DT must include bidirectional communication instead of having a virtual model that only updates according to a physical system." This distinction between the types of communication and interaction is critical, as it defines the extent to which a DT can be leveraged for control, simulation, and predictive purposes.

Table 2.1 outlines different interaction levels within DT systems, ranging from no interaction, where the physical and virtual systems are disconnected, to unidirectional data flow, where data is transferred from the physical to the virtual model, and culminating in bidirectional communication. The latter represents a fully functional DT, where continuous exchange of data allows the virtual model to influence and control the physical system, achieving real-time synchronization and interaction.

Table 2.1: Levels of Interaction in a Digital Twin system between the physical model and its digital counterpart, adapted from [35]

Level of Interaction	Description
No interaction	Virtual model and physical system are not connected through a network. The virtual model only simulates and models a physical system without any real-time updates.
Unidirectional	The physical system feeds sensor data to the virtual model through a network. The virtual model utilizes data to update the current state and predict future states.
Bidirectional	Both the physical system and the virtual model can send data to each other. The virtual model updates using physical data while the physical system can be controlled through data sent by the virtual model.

2.6 Human-Robot Collaboration in Industrial Applications

Following the detailed exploration of DT and MR, this section discusses how these technologies integrate into HRC, demonstrating significant improvements in interaction, safety and efficacy in practical examples from industry solutions with specific emphasis on remote collaboration.

The articles cited below often reference AR, however, it is important to note that in many cases, the features described under AR align more closely with the concept of MR as a medium for collaboration, as explained in 2.4.3. This distinction is crucial, as the developed system is more aligned with MR's capability to further facilitate shared virtual and physical environments, enhancing interaction and collaborative processes across industrial applications.

2.6.1 Enhancing Human-Robot Collaboration through Augmented Reality and Digital Twin Implementation

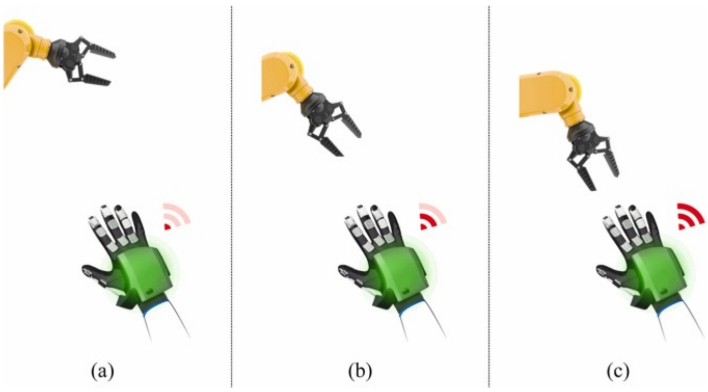
Chu et al. [36] reviewed various studies on integrating AR with DT to improve HRC using visual and haptic feedback interfaces. Their work emphasizes the value of AR in enhancing human-robot communication by providing both egocentric (shared remote views) and exocentric (spatial visualization of the robot relative to the workspace) perspectives, thereby improving spatial awareness and interaction quality.

Green et al. [37] further highlight that multimodal AR interfaces—incorporating visual, haptic, and acoustic cues—can significantly enhance HRC. Multimodal approaches help overcome challenges like limited Field of View (FOV) in HMDs, making the interaction more intuitive. For instance, audio-tactile feedback is particularly beneficial for individuals with visual impairments, providing alternative sensory channels without compromising performance. Visual AR cues, implemented through HMDs, help users navigate complex environments by overlaying relevant information, thus enhancing navigation without impeding robotic movement.

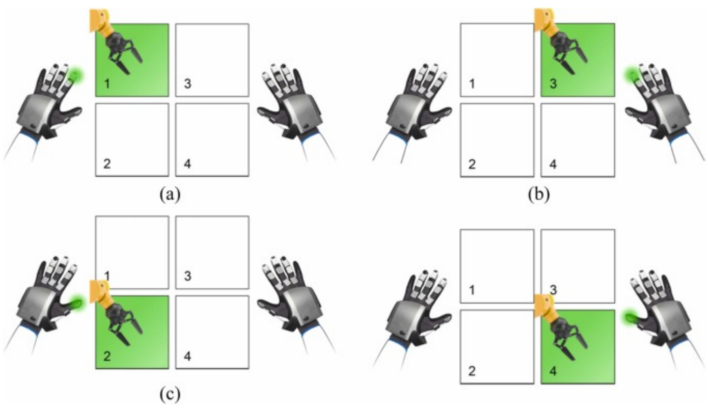
Lasota et al. [38] conducted experiments to evaluate the impact of human-aware motion planning on HRC, demonstrating substantial improvements in task performance and team fluency. Compared to standard robotic systems, participants working with human-aware robots completed tasks more efficiently, exhibited greater concurrent motion, and experienced less idle time for both human and robot. Moreover, they maintained greater separation distances, which reduced collision risks and increased perceived safety. These results illustrate the dual advantage of human-aware planning: it not only enhances task efficiency but also elevates worker comfort and safety, which are critical for minimizing stress-related risks in industrial environments.

The study utilized the ROS to control a WidowX 250 Robot Arm, using the MoveIt framework for motion planning. This approach ensures modularity, adaptability, and ease of management in shared HRC environments. Both AR and DT models were developed using the Microsoft HoloLens 2 for visual feedback and the SenseGlove Nova™ for haptic feedback, offering a comprehensive multimodal experience.

Through the HoloLens, users could visualize the robot's planned trajectory and swept volume, thus anticipating its actions. Concurrently, the haptic interface provided vibration feedback to signal the robot's proximity and target destinations, as illustrated in Figure 2.10. These multimodal cues offered varying levels of detail, aiding coordination in tasks that required awareness of robot movements.



(a) Cues indicating the gripper's destination using vibration on different human fingers.



(b) Cues indicating the proximity of the gripper via vibration frequency changes.

Figure 2.10: Visual and haptic interfaces used in the experiment [36].

Findings show that combining visual, haptic, and acoustic cues significantly improve task performance. Visual interfaces, especially those indicating proximity, excelled in usability, while haptic feedback proved invaluable in scenarios where visual input was insufficient or overloaded. Acoustic signals also served as alerts for sudden changes in the robot's motion, helping reduce operator anxiety in unpredictable environments.

In a task where the operator and robot worked independently but in close proximity, the system enabled efficient coordination. The robot delivered materials while the operator performed assembly tasks, highlighting the potential for AR-based interfaces to optimize HRC in industrial settings.

2.6.2 Augmented Reality-Assisted Multi-Robot Systems for Enhanced Control and Coordination

Integrating AR into multi-robot manufacturing systems offers significant improvements in interaction, operational safety, and efficiency, especially when applied to real-time and planned control modes. Ong et al. [39] further explored AR-assisted robot programming for welding applications, demonstrating that user-friendly interfaces can significantly reduce the complexity and duration of the programming process. These interfaces enable operators to define welding points and orientations using handheld pointers, thus enhancing task accuracy and efficiency by allowing validation within the actual robot workspace.

Malí et al. [40] developed an AR application that permits users to adjust robot axis values, visualize specific robot points through 3D arrows, and navigate hidden points using leading lines. Evaluated in an industrial setting, this application showed improvements in usability and interaction capabilities.

Puljiz et al. [41, 42] explored various AR-based methods for robotic arm programming using devices like the Microsoft HoloLens, implementing techniques such as hand-guided task programming, augmented trajectory visualization, and the creation of spatial maps for virtual waypoint placement. These methods facilitate intuitive and accurate robotic arm programming, enabling seamless integration between virtual commands and real-world operations.

Modern manufacturing trends are characterized by a shift toward mass customization and increased flexibility, driven by the demand for individualized products. This necessitates more adaptable manufacturing systems where human operators collaborate with industrial robots to handle complex tasks [43–45]. However, existing robotic systems primarily execute pre-programmed tasks with limited intelligence.

To address this, two promising approaches have emerged: leveraging advanced AI techniques for robot learning [46] and integrating a human-in-the-loop strategy for robot teleoperation. The latter, more aligned with Industry 5.0 principles, extends the capabilities of both humans and robots by incorporating human expertise into collaborative multi-robot processes [47].

Unlike traditional HRC, multi-robot manufacturing with a human in the loop allows operators to interact with robots from remote locations, not limited to physical workspaces. This paradigm facilitates safer and more flexible manufacturing, by bridging the gap between fully automated and manual operations [47]. However, significant challenges remain, including the need for more user-friendly teleoperation interfaces and systems that can be easily utilized by manufacturing operators without extensive robotics training [48].

Research in multi-agent collaborative manufacturing has focused on enhancing safety, productivity, and cost reduction. Wearable AR-assisted systems and DT technologies enable accurate and intuitive robot teleoperation. By combining AR with robot teleoperation, workers can access physical and virtual information simultaneously in a hybrid environment, interacting with virtual objects [49, 50]. An example is an AR-based teleoperation system utilizing RGB-D imaging, allowing operators to perceive the remote robot's environment and perform teleoperation [51]. Another system transforms robot workspaces into AR environments for rapid and intuitive path planning and task programming [52]. These systems improve task performance by providing additional visual cues to enhance the operator's situational awareness.

Recent advancements in smart manufacturing have led to the development of DT models for robot control. For example, [53] used the Unity engine to create a DT of a robot arm that could learn manufacturing tasks virtually and replicate them in the physical world. The integration of DT and VR interfaces has also been proposed to design immersive human-in-the-loop robotic systems, where the DT acts as an intermediary layer for task execution and quality monitoring [54, 55].

Li et al. [56] demonstrated how AR-assisted **DTs!** enable operators to manage and coordinate multiple robots more effectively. Figure 2.11 depicts an immersive dual view where users can interact with the physical setup of two collaborating robots while simultaneously observing their virtual counterparts through Microsoft HoloLens AR glasses. This setup allows for real-time monitoring and simulation of manufacturing processes, improving robot operation control.

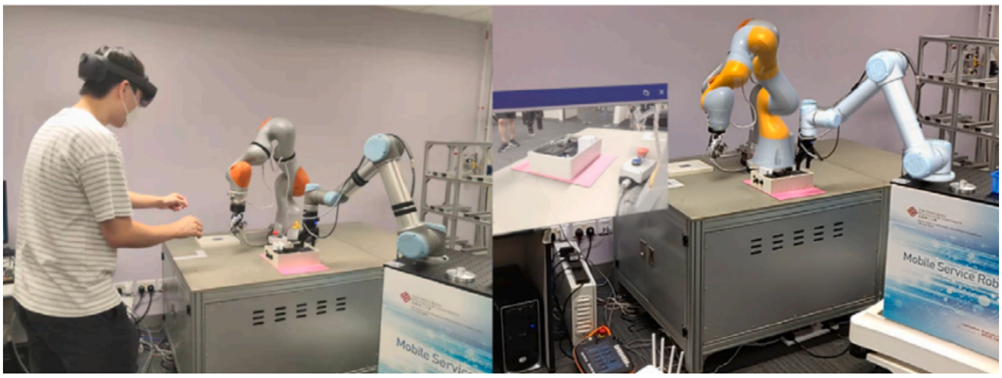


Figure 2.11: AR-assisted DT-enabled multi-robot collaborative manufacturing system [56].

Their proposed comprehensive framework for an AR-assisted, DT-enabled robot collaborative manufacturing system features human-in-the-loop control. The system architecture, shown in Figure 2.12, introduces a multi-node communication mechanism to facilitate interactions among multiple robots and clients. It includes the design of an AR-based teleoperation system for pose registration and motion planning, coupled with three DT-enabled interaction approaches to achieve closed-loop interaction between virtual and physical robots.

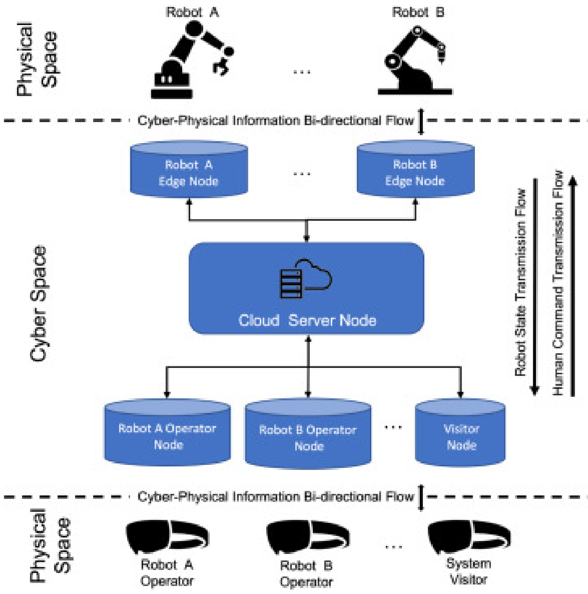


Figure 2.12: The architecture of a multi-robot, multi-client communication mechanism [56].

The DT of the physical robot, modeled using the Unity engine, is displayed as a hologram in the remote workspace via AR glasses, enabling teleoperation and remote monitoring. Pose registration involves aligning the virtual and physical robot models using the Vuforia Engine, while joint alignment translates DT joint values into real-world coordinates.

Therefore, a robot control approach aided by AR technology offers several benefits:

- Enhanced predictability of robot posture and motion trajectories.
- Trajectory visualization to prevent safety issues.
- An intuitive interface that overcomes spatial and physical limitations.

However, observing workspace and robot state information during task execution presents challenges, such as networking latency and positioning accuracy. Proposed solutions include time-sensitive networks and advanced communication technologies such as 5G [56].

The proposed system also utilizes IP cameras for workspace monitoring, projecting video feeds onto AR glasses for enhanced remote monitoring, as shown in Figure 2.13.

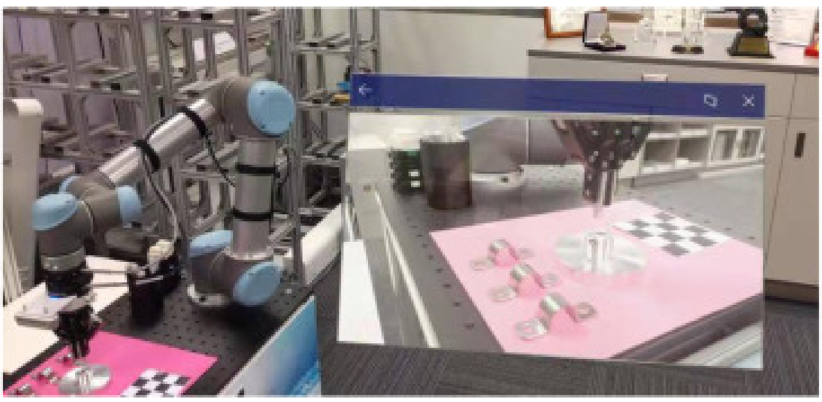


Figure 2.13: Demonstration of the workspace observation approach [56].

2.7 Future Trends in Human-Robot Collaboration

All in all, future directions in HRC are evolving due to advancements in cobot technologies, sensing methodologies, and algorithmic developments [13]. The key trends identified include:

- **Enhanced Scene Understanding:** Next-generation HRC systems will prioritize deeper contextual awareness of the workspace and tasks at hand. This involves not only detecting the physical environment but also interpreting operator intentions, recognizing task progression, and continuously monitoring environmental dynamics. Such enhanced scene understanding will enable robots to anticipate human actions, predict potential safety risks, and adjust their behavior accordingly, thus fostering a higher level of operational safety and efficiency.
- **Advanced Sensing and Data Fusion:** To facilitate this enhanced scene understanding, advanced sensing methodologies and sophisticated data fusion techniques will be critical. By integrating multi-modal sensor data—such as visual, tactile, and auditory inputs—robots will be able to construct more comprehensive models of their surroundings and human collaborators. Real-time fusion of such data will allow systems to process information more effectively, ensuring safer interactions by predicting hazardous movements and improving overall system transparency. This, in turn, will enhance user trust and accelerate the adoption of HRC solutions across industries.
- **Improved Task Planning and Adaptive Learning:** Future HRC systems will be distinguished by advanced task planning capabilities, driven by more sophisticated task modeling and real-time adaptation mechanisms. As robots become more capable of autonomously learning from both structured and unstructured environments, their ability to handle a wider array of tasks will expand, reducing human involvement in routine planning stages. The deployment of these capabilities in manufacturing and service sectors will enable robots to shift between tasks seamlessly, dynamically adjusting their behavior to respond to real-time changes in production or workflow.
- **User-Friendly Interfaces and Interaction Methods:** As the complexity of HRC systems increases, the need for intuitive and accessible human-robot interfaces will become paramount. Developing user interfaces that enable seamless human control without requiring advanced technical expertise is a key area of research. Implementing AR and VR technologies is expected to play a pivotal role in this domain, offering operators immersive and intuitive control mechanisms. These interfaces will reduce cognitive load and enable operators to interact with robots more effectively, thus improving operational efficiency and overall system usability.
- Among other relevant topics, which will not be further described due to not being the focus of this dissertation, incorporating ML and adaptive learning algorithms into HRC systems also represent a transformative leap forward.

Historically, the focus has been on increasing the relevance of HRI by addressing higher safety requirements and enabling robots to perform more complex tasks. Recently, the scope has expanded to include more sophisticated methods aimed at enhancing system performance, applying these methods across different application fields and tackling more intricate tasks. This expansion is driven by the emergence of new cobots, advancements in sensing technologies, matured algorithms, and accumulated experience in designing collaborative workcells [13].

2.8 Summary

Even though significant advancements in AR-DT implementations over the past decade, the state-of-the-art literature predominantly focuses on developing applications for on-site personnel. Remote collaboration, particularly in HRC scenarios, has received comparatively less attention, highlighting the need for systems that effectively facilitate both on-site and remote collaboration. Therefore, the proposed project aims to facilitate and integrate better remote collaboration by proposing a generalized conceptual system applicable across various application scenarios.

Despite having developed on-site features, such as DT pose registration alongside audio and visual cues, focus will be mainly on the remote collaboration part implementation and the billateral communication between users. Unity 3D engine will be further explored for robot model development, ROS will be used for robot control, and Vuforia for pose registration. It will also incorporate visual and audio cues to enhance user safety and awareness, as well as MR elements implemented with Unity 3D. A camera will enable workspace monitoring.

In conclusion, the proposed system will enable remote users to manipulate the robot using Handheld Device (HHD), with the robot's real-time position displayed in the Unity DT. By addressing these challenges, the project aims to enhance remote collaboration in HRC scenarios, contributing to the broader field of MR-DT applications.

*TODO: check if below references are useful

CHAPTER 3

Implementation Tools

In order to integrate Human-Robot Collaboration and Mixed-Reality technologies, a strategic selection of software and hardware tools was essential. This chapter outlines the key technologies and platforms employed to build the conceptual Mixed-Reality based Digital-Twin framework, which facilitates real-time collaboration and bi-directional robot control.

Achieving the successful integration of HRC and MR technologies requires a careful selection of advanced tools, spanning both software and hardware domains. The goal of this dissertation is to construct a robust MR-based DT framework that allows real-time, remote collaboration and bi-directional robot control. The implementation of such a system introduces several key challenges, including the alignment of physical and digital entities, real-time data communication, and the creation of immersive, user-friendly interfaces. This chapter systematically introduces the key implementation tools used throughout the development of the project, explaining their roles in addressing these challenges. The tools discussed in this chapter were chosen for their specific contributions to the project's overall goal.

3.1 UR10e Robot

In order to start adressing the aforementioned challenges, a first effort has been made. A robotic arm from Universal Robots, UR10e, shown in the figure 3.1, available at IRIS LAB.

The UR10e model is one of Universal Robots' most advanced cobots, featuring a payload capacity of 12.5 kg, and a reach of 1300 mm, being designed to automate a wide range of tasks that typically require human input, such as assembly, packaging, and pick-and-place operations [1].

UR10e's integrated force sensors and collision detection technologies allow to collaborate safely with humans in shared workspaces, making it ideal for HRC scenarios. Besides, it offers significant flexibility in terms of programming and adaptability. Its console is intuitive

[1] UR10e `https://www.universal-robots.com/products/ur10-robot/` Accessed: 2024-10-15

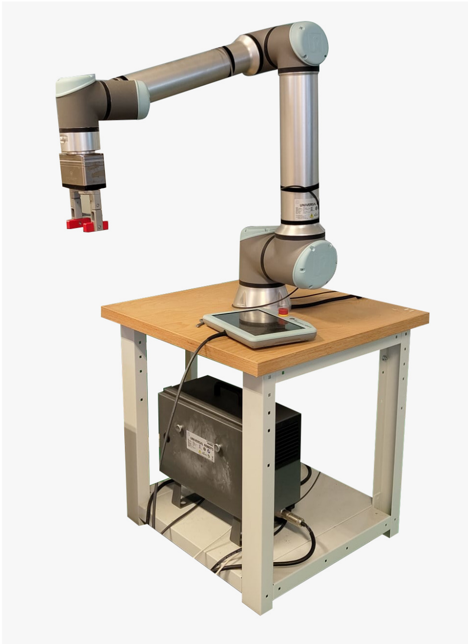


Figure 3.1: UR10e Robot used in the IRIS-LAB, University of Aveiro

and allows imported pre-programmed scripts, therefore being easily deployed across various industrial tasks with minimal programming experience required by the operator.

This robot, as the core physical entity in this human-robot collaborative system, serves as the dynamic agent for performing collaborative tasks, where its physical attributes were mirrored in an immersive digital environment, regarding the fundamental DT concept.

3.2 Simulation Environment

3.2.1 Unity

Unity, developed by Unity Technologies, was selected as the primary platform for developing the MR environment in this project. Originally designed for game development, Unity has evolved into a powerful tool for creating interactive 3D applications, including AR, VR, and MR. Its robust architecture and versatility in rendering complex virtual environments make it an ideal choice for building a dynamic DT of the UR10e robotic arm. This ability to seamlessly integrate external data sources, such as sensor inputs from real-world hardware, enables a high degree of interactivity and realism in the simulation.

Unity's Integrated Development Environment (IDE) allows for rapid prototyping and iterative design of both the virtual space and the MR user interface. Moreover, the engine's cross-platform compatibility supports a range of devices, including desktop systems and mobile platforms, and can handle real-time rendering of high-fidelity 3D models, essential for MR applications. Through the Unity-based simulation, operators can not only control the physical UR10e robot remotely but also visualize real-world tasks in a digital, augmented environment, ensuring accurate synchronization between physical and virtual realms.

In addition, Unity's robust asset management and scripting support, primarily through C#, provide developers with the tools to easily simulate complex environments, manage interactive objects, and implement advanced functionality such as collision detection and user input handling. These features enable a realistic and immersive experience for both on-site and remote users, improving the overall effectiveness of the HRC system.

3.3 Digital Model Implementation of the Robot

3.3.1 Digital Robot Model - URDF Importer Package

In order to correctly mirror the UR10e physical robot model and establish a functional DT bidirectional synchronization, the Unified Robot Description Format (URDF) model of the robot was imported to Unity simulation environment. Unity Robotics Hub's URDF Importer package was used for this very purpose [2]. This package provided several key advantages during the implementation of the DT for the UR10e robot.

The *Unity URDF Importer* package significantly facilitated the integration of the URDF file, enabling the precise recreation of the UR10e robot's physical structure, including its joints and linkages. This seamless integration supported real-time simulation of the robot's movements and configurations within the Unity environment. Additionally, it allowed the visualization and fine-tuning of physics properties to accurately reflect the robot's real-world dynamics. Essential for bilateral communication between virtual and physical robots, the package also streamlined sensor data import from ROS, enhancing the realism of simulations and optimizing development and testing processes.

3.4 Pose Registration

Pose registration is a crucial step in aligning the digital model with the physical robot. In this project, Vuforia, a cutting-edge AR software platform, was integrated with Unity to accomplish this alignment.

3.4.1 Vuforia

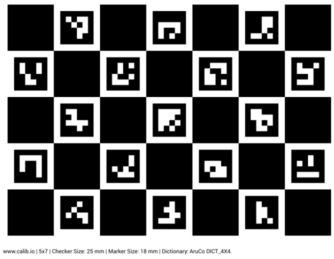
Vuforia Engine Software Development Kit (SDK), an advanced AR development platform that can be used within Unity platform, provides robust capabilities for object recognition and tracking, making it integral to MR applications. In this project, Vuforia's image-based tracking system was employed to align the digital model of the robot with its physical counterpart. This was achieved through the detection and tracking of an ArUco marker placed next to the physical robot, ensuring precise spatial alignment between the digital and physical elements within the MR environment, thus enhancing real-time interaction accuracy.

3.4.2 Marker Detection

An ArUco marker, illustrated in the Figure 3.2a, served as a medium to perform accurate pose estimation of the robot within the Unity environment. Besides, the Logitech c922 camera

[2]Unity Robotics Hub `https://github.com/Unity-Technologies/Unity-Robotics-Hub` Accessed: 2024-02-02

shown in Figure 3.2b scanned this marker, enabling the system to overlay the digital model accurately over the physical robot. This ensured precise positioning and manipulation of the digital twin in the MR space.



(a) ArUco marker for alignment of the digital twin

(b) Logitech c922 camera used for marker detection

Figure 3.2: Marker and camera setup for digital and physical robot alignment

In Figure 3.3 the digital robot model can be seen positioned realitve to the above described marker.

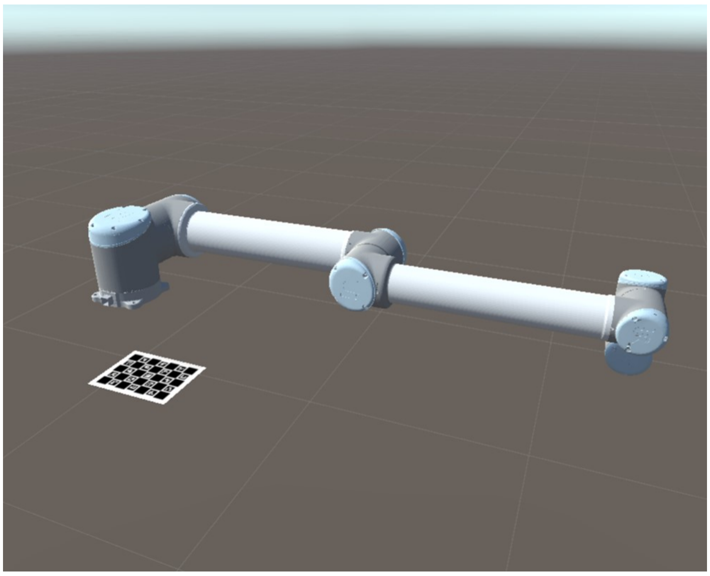


Figure 3.3: Digital UR10 model aligned with ArUco marker in Unity

* TODO: Add a figure showing the digital UR10 model aligned with the real ArUco marker and the real robot in the background. This will visually represent the alignment process - this figure should be displayed here or in the 4 chapter

3.5 Bidirectional Communication

After having the digital model correctly aligned with the physical robot, the next step was to establish bidirectional communication between the Unity and the UR10e. This communication was essential for enabling remote control of the physical robot through the DT, as well as for synchronizing the robot's state between the real and virtual environments.

3.5.1 ROS

The Robot Operating System (ROS) was chosen as the middleware for facilitating real-time communication between the physical robot and the Unity environment, since the UR10e robot from IRIS-LAB already had prior developed ROS packages, such as `iris_ur10e` [3] and `iris_sami` [4].

These packages provided a comprehensive ROS setup for controlling the UR10e robot, including trajectory planning, RViz visualization, and real-time operation. ROS offers a robust and flexible framework for developing complex robotic systems. In this project, it enables seamless integration between hardware components and software modules. Specifically, ROS facilitates the exchange of sensor data, control commands, and state information between the physical UR10e robot and its digital twin in Unity, ensuring synchronization for remote control and collaborative operations.

[3] IRIS-LAB github Repository `https://github.com/iris-ua/iris_ur10e`, Acessed: 2024-02-02

[4] IRIS-LAB github Repository for UR10e Robot Manipulation `https://github.com/iris-ua/iris_ur10e`, Acessed: 2024-02-02

CHAPTER 4

Mixed Reality for Human-Robot Collaboration

This chapter presents the development of a comprehensive framework for implementing a Mixed Reality environment in Human-Robot Collaboration. The framework leverages advanced technologies such as Mixed-Reality, Digital Twins, and ROS, ensuring precise synchronization between physical and digital models. It discusses the tools, control methods, and User Interface enhancements designed to optimize both on-site and remote user experiences, aiming towards an intuitive, immersive collaboration.

4.1 Framework

Figure 4.1 illustrates the proposed framework of the MR-based HRC system. This framework seamlessly integrates two distinct environments—on-site and remote—through a robust communication pipeline designed to enable real-time, collaborative robot manipulation.

First of all, regarding the **On-Site Environment**, the UR10e robotic arm operates as the central physical entity to be controlled and manipulated. The on-site user interacts with this robotic arm through a custom MR application developed in Unity, chosen due to its robust MR capabilities, which facilitate the creation of immersive, interactive environments, enabling intuitive real-time robot manipulation.

After implementing the robot digital model into the simulation environment, the next step was to align it with the physical robot. Pose registration between the physical robot and its digital counterpart is executed using Vuforia's capabilities. The system employs ArUco markers for precise pose estimation, ensuring that the digital representation of the UR10e is accurately aligned with its physical counterpart. The DT of the robot, rendered in Unity, provides a visually synchronized, real-time mirror of the robot's movements and configurations, thus facilitating enhanced interaction.

The robot is connected via Ethernet to a laptop running Ubuntu 20.04 with ROS Noetic, which serves as the middleware layer. This setup facilitates seamless data exchange between the

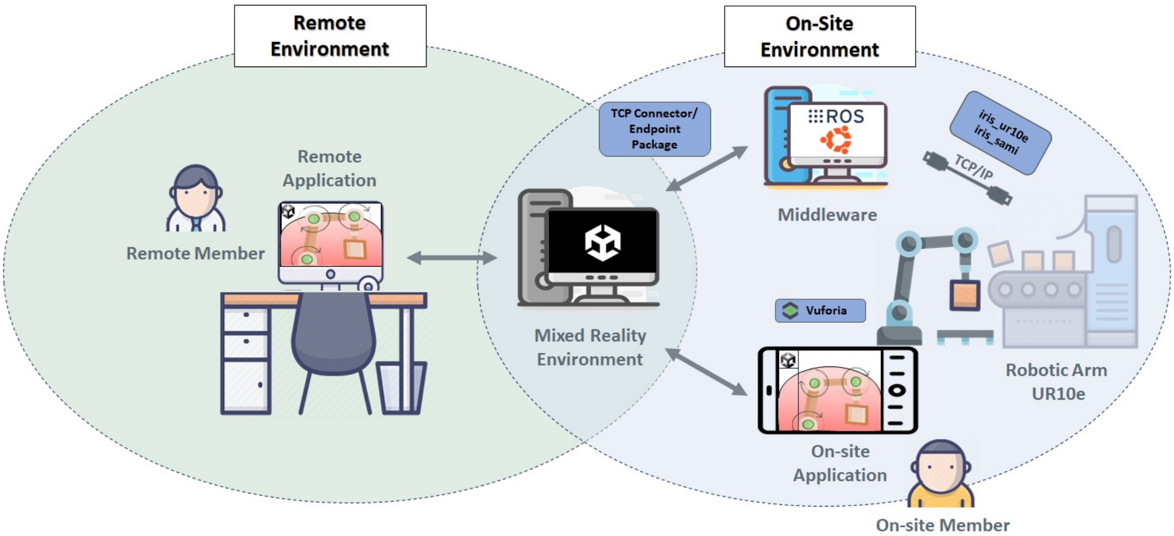


Figure 4.1: Overview of the proposed MR-based HRC system framework integrating remote and on-site environments.

Unity DT and the physical robot. The `iris_ur10e` and `iris_sami` ROS packages, developed by the IRIS Lab and available on GitHub[1], provide a pre-established ROS environment that supports critical functionalities such as trajectory planning and robotic manipulation and visualization through RViz.

To tailor these packages to the specific needs of this project, several enhancements were made. These modifications included the integration of bidirectional data flow between ROS and Unity, enabling the Unity-based DT to mirror the real-time movements of the physical robot. In particular, new ROS nodes were created to subscribe to joint state data from the physical robot and publish these updates to Unity, ensuring precise synchronization between the physical and virtual environments. Additionally, new publishing mechanisms were implemented to send commands from Unity back to ROS, allowing for full control of the robot through the MR interface. These modifications were essential for realizing the bidirectional communication required for accurate DT manipulation.

This communication between the ROS middleware and Unity's MR environment is established using Unity's ROS-Transmission Control Protocol (TCP)-Connector and ROS-TCP-Endpoint packages. These packages enable bidirectional communication over a TCP/Internet Protocol (IP) protocol, ensuring real-time synchronization between the virtual and physical environments. This communication architecture is fundamental for maintaining the DT's fidelity, reflecting real-world changes in the Unity model and vice versa, as referred in the section 2.5.

Regarding the **Remote Environment**, a remote participant accesses the same Unity-MR application. This allows the remote member to visualize and manipulate the robot in real time, from a separate location. The remote UI provides real-time visualization of the robot's state and its workspace, enabling remote collaboration. The synchronization between the

[1]https://github.com/iris-ua

remote and on-site environments is facilitated through Unity's MR capabilities, which, in conjunction with the ROS-based control, enable the remote user to execute commands and receive real-time feedback.

The Middleware layer, acting as the system's backbone, ensures the continuous synchronization of data between the physical robot and its DT. It manages the real-time feedback loop, maintaining bidirectional data flow between the virtual robot in Unity and the physical robot in the on-site environment. This configuration guarantees that any actions performed by either the on-site or remote user are consistently reflected in both the physical and digital realms, preserving operational coherence and maximizing collaborative efficiency.

This framework provides an immersive and responsive MR environment, bridging the gap between physical and digital spaces. The system enables real-time robot manipulation and monitoring from both on-site and remote locations, making it a versatile platform for collaborative tasks in advanced industrial applications. The seamless integration of MR, DT, and HRC technologies significantly enhances user interaction, safety, and productivity, while offering an intuitive interface for remote and on-site collaboration.

Three distinct control methods were developed within the application to facilitate manipulation, which will be explained further below.

4.2 Mixed Reality Environment

Regarding the Unity developed MR environment, it started by implementing the proper UR10e digital model. This model was then aligned with the physical robot through Vuforia's pose registration capabilities, ensuring that the DT would accurately mirror the robot movements and configurations.

The base robot manipulation provided by the URDF-Importer package consisted on moving each joint individually by selecting the desired joint and then moving it with by using the keyboard directional arrows. This method was improved to three distinct control methods that enhance user experience and facilitate robot manipulation.

4.3 Robot Manipulation

In order to properly develop new ways of controlling the DT version of the robot, it was necessary to understand the C# script named `Controller.cs`, which contained all the necessary functions to control the robot's joints.

After further analysis, three methods were developed to control the robot's joints:

4.3.1 UI Control

This method is designed to provide users with an intuitive, user-friendly interface for manipulating the DT of the UR10e robot.

Figure 4.2 illustrates the UI developed in Unity, highlighting its primary control panel on the left (1), which includes a detailed joint selection menu, directional arrows, and a reference image of the robot with joint labels (2 and 3). This panel is central to controlling the digital model of the robot, while other UI features shown in the figure will be explained in subsequent sections.

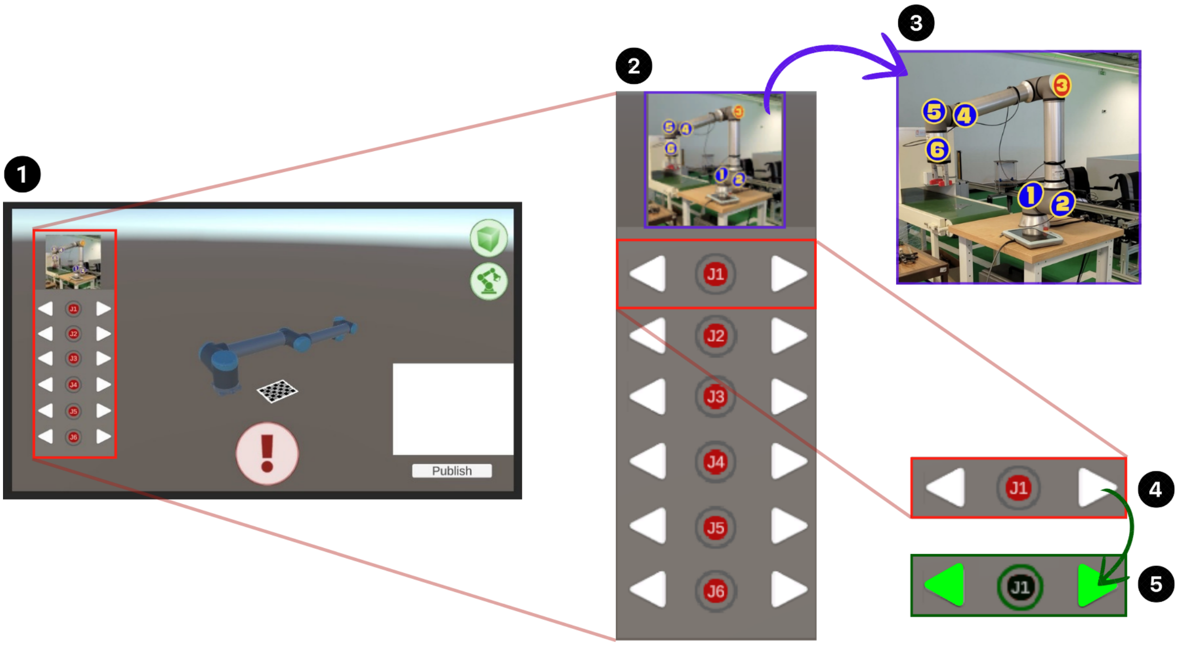


Figure 4.2: UI panel (2) for the MR Unity environment, enabling control of the digital model of the UR10e robot using the UI Control method. Users can manipulate each joint individually, with active joints displayed in green to indicate selection (5). The reference image of the robot (3), labeled with joint numbers (J1 to J6), aids in joint identification from base to end-effector. Directional arrows for each joint (2,4,5) facilitate movement control in positive or negative directions, offering intuitive joint-level manipulation within the simulation.

Interface Structure and Joint Manipulation:

- **Joint Selection**: Each joint has a central button labeled with its identifier (e.g., J1, J2). To activate a joint, the user clicks the corresponding button, which changes color from red to green, signaling that the joint is selected for movement (4 and 5). This selection process helps avoid confusion and accidental manipulation of multiple joints.
- **Movement Control**: Once a joint is selected, the user can manipulate its position by clicking the directional arrows on either side of the central button. Pressing the right arrow rotates the joint in the positive direction, while the left arrow rotates it in the negative direction. This control mirrors the real-time responsiveness typically provided by keyboard arrows, ensuring an intuitive experience.
- **Continuous Movement**: The selected joint will continue moving in the chosen direction as long as its central control button and the directional arrow remain selected. To stop the movement, the user can simply deselect the joint by clicking the central button again, which reverts to its default color, or deactivate its rotation direction.

Operational Rules:

- **Single Joint Activation**: Only one joint can be active for rotation at any time. If multiple joints are selected (i.e., their buttons are green), the system prevents movement in any direction until only one joint is selected.

Additional Design Considerations:

- **Visual Reference for Joint Positioning**: The overlay image at the top of the joint control panel serves as a quick-reference guide for users to confirm joint positions relative to the physical robot. This visual aid is useful in remote collaboration or complex tasks where clear identification of each joint's location is essential.
- **UI Responsiveness and Feedback**: To ensure a smooth interaction, the interface provides visual feedback for each action (e.g., button color changes) and restricts movement based on user inputs, making the control process both transparent and manageable.

This structured approach to joint manipulation promotes safe, precise, and efficient interaction with the robot, supporting the usability goals of the MR application. By minimizing the likelihood of accidental commands and providing continuous, clear visual feedback, the interface helps maintain control clarity and operational awareness throughout the interaction.

4.3.2 Unity to Robot

Similarly to the previous described method, this one also enables the digital robot model to be controled from the MR environment. However, when needed, it updates the newly defined digital robot position into the real on-site UR10e robot. In order to change the DT, the user has to press the right/left arrow keyboard keys to select the following/previous joint as well as the up/down keys to rotate the selected joint in the positive/negative direction, respectively.

After moving the digital model to the desired position, by pressing the "Publish" button within the UI, as shown in Figure 4.3, the robot's joints coordinates are published to the ROS middleware, using the ROS-TCP-Connector/Endpoint packages.

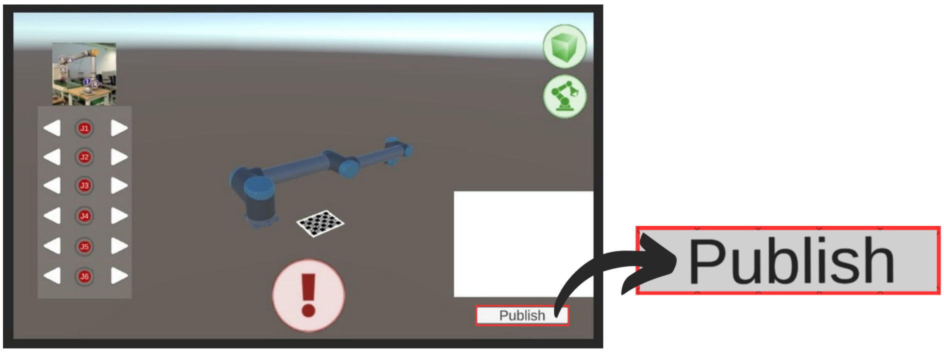


Figure 4.3: Publish button that sends Unity's DT robot joint states into ROS the environment

To avoid conflicts with the ROS node responsible for controlling the real robot's joints, which are being constantly published to the standard `joint_states` topic, a separate ROS topic (`unity_joint_states`) was created to handle the joint data coming from Unity. This ensures that data from Unity does not interfere with the real robot's ongoing operations. When the "Publish" button is pressed, the Unity-defined joint states are sent to this new topic, and only when necessary are they relayed to the real robot for movement execution. Below, there is a pseudo-code snippet explaining how to update the DT in Unity and send the new desired robot position to the ROS middleware.

Algorithm 1 Unity Input for Joint Selection and Movement

```
Step 1: User Input for Joint Selection and Movement in Unity
while Unity Simulation is running AND Unity-ROS Control is enabled do
    if RightArrowKeyPressed then
        Select next joint
    else if LeftArrowKeyPressed then
        Select previous joint
    end if
    if UpArrowKeyPressed then
        Rotate selected joint in positive direction
    else if DownArrowKeyPressed then
        Rotate selected joint in negative direction
    end if
    if PublishButtonPressed then
        joint_states = GetCurrentJointStates()
        PublishToROSTopic('unity_joint_states', joint_states)
    end if
end while
```

In order to handle this communication process, two new ROS nodes were created, the `joint_state_listener` and the `move_unity` nodes. The `unity_joint_subscriber.py` script was developed within the `iris_ur10e` package, initializing the node that subscribes to the `unity_joint_states` topic, which receives a JointState message type. It then initializes the publisher for the `move_joint_unity` topic, converting this data into a Float64MultiArray format which will be further received by the second node. This second node called `move_unity`, created within the `iris_sami` package, subscribes to the `move_joint_unity` topic, listening to new joint position values, moving the robotic arm to this desired position. This real-time update can be visualized either on the simulation environment, through Rviz, or by utilizing the real UR10e robot. Another pseudo-code explanation regarding how both ROS nodes work, is presented below.

Algorithm 2 Combined ROS Node for Receiving Unity Joint States and Moving the Robot

Step 2 and 3: Combined ROS Node for Receiving Unity Joint States and Moving the Robot
Initialize ROS Node: joint_state_listener
Subscribe to Topic: 'unity_joint_states'
while Receiving JointState message from Unity **do**
    float_array_data = ConvertToFloat64MultiArray(joint_states)
    PublishToROSTopic('move_joint_unity', float_array_data)
end while
Precondition: The `joint_state_listener` node must be running and publishing to the `'move_joint_unity'` topic.
Initialize ROS Node: test_arm_movement
Subscribe to Topic: 'move_joint_unity'
while Receiving Float64MultiArray message from `move_joint_unity` **do**
    MoveRobotArmTo(joint_positions)
    if ConnectedToRealRobot **then**
        MoveRealRobot()
    else
        VisualizeInRviz()
    end if
end while

4.3.3 Robot to Unity

* TODO: add here a figure of the Robot to Unity - real robot and superimposed dt in unity

Opposite to the above described method which controlls the DT and then updates the robot position, in this control method, the on-site user moves the robot and the remote counterpart visualizes this update instantly within the MR environment.

In order to properly achieve this communication and data transfer, the `JointStateSubscriber.cs` script was created in Unity. It subscribes to the `joint_states` topic, and stores the information regarding the joint positions in a dictionary structure that is updated in real time into a specific `.json` file within the MR environment. This file is constantly being read by the `Controller.cs` script whenever this control method is enabled, updating the DT robot model accordingly.

By maintaining this synchronization between the real robot and the virtual environment, the Unity scene accurately reflects the robot's live state, ensuring a consistent DT representation through the bidirectional communication established between the MR environment and the ROS middleware. Below, there is another pseudo-code explaining how the ROS-Unity control method works.

Algorithm 3 ROS-Unity Control via Joint States Subscription

Step 1: Subscribe to ROS `joint_states` topic
Attach the Unity Script to the Digital Robot Model Asset: `JointStateSubscriber.cs`
Upon Initialization, it subscribes to topic: `/joint_states`
while Receiving JointState message from ROS **do**
    Extract joint names and positions from the message and store them in a dictionary data structure
    Save the dictionary data to the `jointStateSubscriber.json`
end while
Step 2: Update Unity DT Robot Model
while Simulation is Running **do**
    Read the `jointStateSubscriber.json`
    Update the Unity DT robot model using the joint positions from the dictionary structure
end while
Step 3: Synchronize Real Robot with DT Robot
The Unity DT robot model moves according to the real robot's joint positions, ensuring a consistent bidirectional DT representation.

4.4 On-site Mixed-Reality Features

After having implemented the DT-Robot bidirectional communication between the on-site and remote members, the next step consisted on developing features that could enhance both users' experience when interacting with the collaboration environment. These features were designed to improve user safety, facilitate robot manipulation, and provide an intuitive interface. The following sections detail these key features development and implementation.

4.4.1 Virtual Safety Zones

As explained in section 2.6, introducing different sensorial cues enhances on-site users' experience into a more intuitive and immersive experience. Therefore, visualizing the working zone of the robot is a critical feature designed to enhance safety when interacting with the robot. In order to achieve this, two safety zones were developed, as shown in Figure 4.4. * TODO: re-do this figure with the most recent version of the UI and safety-zones

Addressing specific safety and user experience concerns:

- **Outer Safety Zone:** Initially, only this safety zone was created. The purpose of creating it was to provide an early warning to users as they approach the hazardous area near the robot. This approach consisted on changing its color as a visual alert. However, this method proved ineffective because, once inside it, users could not perceive the color change, rendering the warning system inadequate.
- **Inner Safety Zone:** To overcome this outer zone limitation, an additional Inner-Safety Zone was developed. This design ensures a two-step safety mechanism that properly alerts users when they are in close proximity to the robot.
- **Sensorial Cues:**

Visual: Upon entering the Outer-Safety Zone, the color of the Inner sphere changes to red, reverting to its default color if the user exits this critical area. This visual cue alerts the user to increased proximity to a high-risk zone. Additionally, a blinking warning appears at the bottom center of the interface, switching between both circular danger signs shown in Figure 4.4. This flashing indicator amplifies the alert, reinforcing the awareness of approaching the operational robot area, remaining active as long as the user is within this Outer-Safety Zone.

Auditory: Entering the Inner-Safety Zone triggers an audio alarm signifying that the user has breached into the robot working area, enhancing the effectiveness of the safety mechanism.

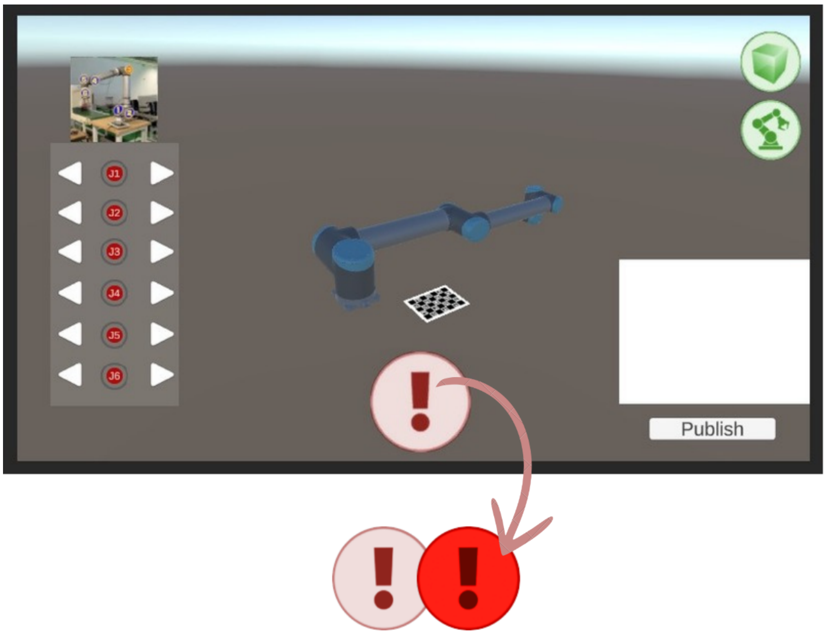


Figure 4.4: Warning blinking sign displayed in the bottom part of the UI to alert the on-site user of proximity to the Robot

- **Breach Protocol:** Besides the above described sensorial cues whose goal consists on improving user's awareness, another feature was implemented to ensure user's safety when interacting with the robot. If the on-site counterpart enters the Outer-Safety Zone area while the robot is in motion, the robot automatically stops. This immediate halt ensures that potential accidents or injuries are avoided by preventing any interaction with the robot when a user is within this designated dangerous area.

4.4.2 Additional Interface Features

Beyond the joint control panel explained in Section 4.3.1, several additional features were integrated into the MR application aimed at enhancing user flexibility and clarity. Positioned in the upper right corner of the UI, two green toggle buttons provide control over distinct interface functionalities.

The lower button, as shown in Figure 4.5, enables users to activate (1) or deactivate (2) the joint control panel, changing to gray when inactive. Deactivating this panel provides an unobstructed view of the shared workspace, enhancing awareness for both on-site and remote users. The upper button operates similarly but controls the visibility and functionality of the safety zones. When toggled off, both safety zones disappear and their associated alerts are disabled. This configuration allows users to customize the interface to match the specific task requirements, reducing visual distractions when safety zones are unnecessary.

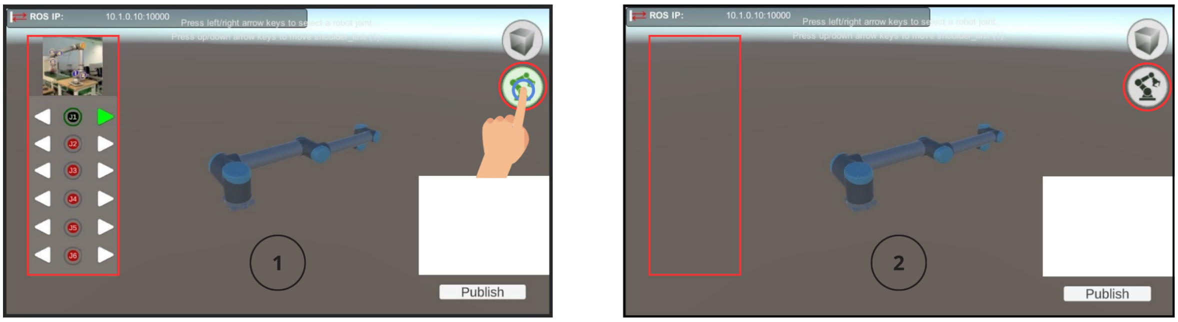


Figure 4.5: Illustration of the activation and deactivation functionality for the robot control panel within the MR interface. In the left image (1), the control panel is fully active, displaying joint selection buttons as well as corresponding directional arrows, along with a reference image of the robot. The right image (2) shows the interface after control panel deactivation by pressing the toggle button (highlighted in red), enabling a clearer view of the environment.

4.4.3 Camera Feed Transmission

To improve situational awareness for remote users, an additional camera was integrated, providing real-time visual context directly from the robot's perspective and reducing on-site operator's responsibility for managing environmental views.

The Orbbec Astra 3D camera [2], shown in Figure **??**, was chosen for its high-resolution video output and compatibility with the ROS framework. This camera was mounted directly onto the robot, offering an immersive view that aligns with the robot's operational perspective. Integration into the ROS environment was facilitated through the Astra camera's dedicated GitHub repository [3], which provided essential drivers and nodes for ROS compatibility.

The live feed from the Astra camera was managed via the `astra_camera_node` from the `astra_camera` package, enabling continuous data capture and real-time display in RViz for preliminary verification. This configuration ensured proper camera functionality within the ROS environment and accurate data capture. However, when transmitting the uncompressed image data to the Unity MR application, bandwidth and latency challenges arose, potentially impacting real-time collaboration. To address this, the ROS `image_transport` package was implemented, which converted the raw video stream into a compressed format, reducing data size while preserving essential image quality. The following command facilitated the data compression:

```
rosrun image_transport republish raw in:=/camera/color/image_raw
```

[2] `https://www.orbbec.com/products/structured-light-camera/astra-series/` Acessed:2024-10-22
[3] `https://github.com/orbbec/ros_astra_camera` Accessed: 2024-10-04

```
out:=/camera/image_repub
```

By compressing the image data, the system achieved more efficient and stable real-time transmission to Unity, providing continuous visual feedback with minimal latency. Afterwards, this compressed camera feed was integrated into Unity using a custom `CameraFeedReceiver.cs` script, responsible for receiving and rendering the feed into the UI panel. This setup allowed the remote user to view the robot's perspective in real time, independent of the on-site operator's actions.

This dual-view capability offers the remote user a comprehensive visual overview by combining the on-site provided view with the camera feed directly from the robot's perspective, potentially enhancing spatial awareness and operational context. Figure **??** demonstrates the live feed display in Unity, showcasing how the system allows remote users to monitor the robot's surroundings and adjust their commands accordingly. This feature may support more informed decision-making in future implementations of remote MR systems for HRC. *TODO: add a figure of the camera transmission in unity and rviz. should i say that it provided situational awareness of robot's surroundings, since it was not atached to the robot itself.

*TODO: is this useful to any integration part of this chapter???? The primary goal of this project consists on enhancing remote collaboration between human operators by utilizing a robotic arm (UR10e) and MR. In order to achieve this, a framework allowing for an interactive, fully functional MR application that facilitates bidirectional communication between the robot and its DT was developed. It features a friendly UI that enables the visualization and control of the robotic arm, where the DT was designed to respond in real-time to the robot's physical movements, and vice versa. The achieved bidirectional communication ensures that the MR system operates not as a mere digital shadow of the robot but as a true DT, capable of reflecting and influencing the physical entity.

One of the key features developed was a seamless control method integrated within the MR environment, allowing the user to manipulate the robot via the UI. This interface includes joint's selection buttons, directional controls, and toggle switches for precise manipulation of each robotic joint. This setup enables users to control the robot's movements through the MR interface and send this position update to the ROS middleware, with real-time synchronization ensuring the robot accurately mirrors the DT's state.

The development of two safety zones within the UI was a critical enhancement aimed at ensuring user safety and improving operational awareness when interacting with the robot. These zones function by providing real-time alerts to the user. When the user enters the "Outer Safety Zone", which has a larger radius, not only a visual alert starts blinking, but also the color of the "Inner Safety Zone" changes, signaling that the user is approaching the robot's workspace. These visual cue escalate the user's awareness of proximity to the robot. If he gets even closer to the robot and breaches the "Inner Safety Zone", an auditory alarm is triggered, continuously alerting the user to their presence within a hazardous area. This layered approach—combining visual color changes and auditory alarms—ensures that users are fully aware of any potential danger during robot operation, thereby preventing accidents.

Moreover, the combination of these sensorial cues also contributes to a more immersive interaction experience, increasing awareness without overwhelming the user. However, the accuracy of these safety zones is heavily dependent on the alignment between the camera and the AR marker. Misalignment between the them can lead to inaccuracies in distance measurements, affecting the reliability of the safety zones. Maintaining precise marker tracking is therefore essential for ensuring both the accuracy of safety zone alerts and the overall safety of the system during operation.

Another significant feature was the implementation of a live camera feed, allowing the remote participant to view the robot's workspace in real-time. This feed is crucial for remote collaboration, enabling users from different locations to have a synchronized understanding of the robot's surroundings. The camera feed is transmitted from the ROS middleware to the Unity MR environment via TCP/IP after image compression. Compression was necessary due to the substantial bandwidth required for raw video data transmission. This feature offers an important perspective for remote users, aiding in monitoring and providing assistance when necessary.

Both the joint control interface and the safety-zone mechanisms were designed to be toggled on/off, thereby ensuring that the user has the option to clear the UI for an unobstructed view of the environment.

CHAPTER 5

Discussion

*TODO: por introducoes em todos os capitulos *TODO: nao ha necessidade de constatar o que foi implementado de novo, porque isso ja esta no capitulo anterior. por o caso de futura aplicacao neste capitulo, 1 ou 2 *TODO: reformulate how does it make sense, since there were no performed user tests, how should i say that no evaluation can be done, since the focus remained on implementing the interface and functionalities and not in perform user test cases - there was time for this

In this section, a critical discussion of the developed MR environment will be done, regarding this dissertation's objectives.

5.1 Mixed Reality reflection on the proposed framework

5.1.1 Challenges and Issues Faced

Though the development of these functionalities was successful, several challenges were encountered during the implementation phase that required substantial effort. The first hurdle was the integration of the robot's digital model into the Unity environment using the Unity Robotics Hub's `URDF Importer` package. Unfortunately, an exact UR10e model was not available on-line, and a UR10 model was used instead [1]. This change did not pose any significant problem, as the physical properties and kinematics between the two models are the same, allowing for accurate simulation.

Ensuring accurate pose registration and alignment between the physical robot and its DT in Unity presented a challenge. This required using a marker for precise AR visualization. After testing several methods, an ArUco marker proved most effective for reliable tracking and better model alignment. Implementing this through Vuforia SDK within Unity, along with relative positioning between the marker and digital model, was essential for maintaining synchronization between the physical and digital entities. However, practical issues arose

[1] PositronicsLab `https://github.com/PositronicsLab/reveal_packages/tree/master/industrial_arm/scenario/models/urdf/ur10` Accessed: 2024-02-05

during testing, as the camera and laptop needed to be moved frequently to maintain proper marker alignment.

When establishing the bidirectional communication between the ROS middleware and the Unity MR environment, the Unity Robotics Hub repositories played a critical role [2]. Integrating these tools was not straightforward due to the lack of comprehensive documentation, particularly around generating and handling ROS messages from within Unity's `C#` environment. ROS messages needed to be generated, transmitted, and handled effectively, but interfacing between `C#` and ROS message types such as `JointState.msg` proved challenging. This complexity arose in part while dealing with the intricacies of ROS message handling in Unity's `C#` ecosystem and the need to ensure that messages were correctly serialized and deserialized for communication. Once the messages were correctly generated on the Unity side, the robot's joint states were transmitted from ROS and saved as a `JSON` file in Unity. This `JSON` file was then used to update the DT in real-time, maintaining the synchronization between the physical robot and its virtual counterpart. The reverse process, where joint coordinates are sent from Unity to ROS, also posed difficulties due to the complexity of publishing these messages back to the ROS environment. This step was crucial to ensure that changes made in the Unity environment could correctly influence the physical robot.

One additional challenge was the live camera feed transmission from ROS to Unity, which proved to be too large to transmit efficiently over the network, resulting in significant latency. To address this, the image data had to be compressed using the `image_transport` package in ROS, which allowed the camera feed to be republished in a more compact format. This enabled real-time transmission over a TCP/IP connection, ensuring that the remote user could observe the robot's workspace without excessive delay.

In conclusion, while the implementation of the core functionalities of the MR application—such as robot control, bidirectional communication, and user interface elements—was successful, the project encountered significant obstacles mainly related to the lack of documentation that enables the integration of external tools and the handling of ROS messages in Unity. Despite these challenges, the system ultimately achieved the goal of creating an immersive and functional MR remote collaboration platform.

*TODO: mention that the control UI can be accessed either in laptop in unity simulation or in an HHD, which was the main purpose, but there were issues when building the unity app.

5.1.2 Use Case Scenario: Collaborative Robot-Assisted Assembly Task

The developed MR system can be applied to a wide range of collaborative tasks involving HRI. For example, in an assembly task, the on-site user and the remote expert can work together using a robotic arm, aided by the system's MR interface. This enables both users to visualize the workspace and the robot's actions in real-time, enhancing their collaborative decision-making.

[2] `https://github.com/Unity-Technologies/Unity-Robotics-Hub`

Core functionalities of the system include using a DT to ensure synchronization between the physical robot and its virtual counterpart, allowing both users to control and monitor the robot in real-time. The live camera feed from the robot provides additional visual context for remote experts, enabling precise instructions and adjustments during complex tasks. Safety zones are implemented to ensure that on-site operators can interact safely with the robot, halting operations if predefined zones are breached.

Application Areas

This developed MR framework can be adapted to industries such as automotive or electronics assembly, as well as healthcare, where precision and real-time collaboration are essential. The remote expert can guide on-site technicians in performing intricate tasks while maintaining full visibility of the robot's actions through the MR interface. However, regarding ergonomics and safety features, these would need to be tailored to industry-specific needs, ensuring that the system promotes both efficiency and user well-being.

For example, a use case scenario could be using the MR system to facilitate collaboration between an on-site technician and a remote expert during a complex LEGO assembly. The on-site user arranges the LEGO pieces while the remote expert guides the process via the MR interface, visualizing the robot's workspace and controlling the UR10e robotic arm. While the on-site user positions larger LEGO blocks, the remote user, viewing the real-time camera feed and DT synchronization, handles intricate placements with precision. Both users interact with the robot using the MR interface, allowing seamless handoffs and clear coordination. Virtual safety zones ensure that on-site users are protected from accidental robot movements. This application highlights how the system can improve teamwork, task accuracy, and efficiency in precise, component-based tasks, with the real-time updates and collaboration capabilities ensuring minimal errors during assembly.

On the other hand, regarding the healthcare setting, the MR system could be deployed to assist in a remote surgical procedure. An on-site surgeon collaborates with a remote expert who oversees the operation via the MR interface. The robotic arm assists with precise movements during surgery, such as handling instruments or holding tissues. Again, the live camera feed from the robot provides the remote expert with a surgeon's view of the operation, while the DT in the MR interface mirrors the robot's real-time movements. The remote expert can adjust the robot's actions, guiding the on-site surgeon through critical parts of the procedure. Gesture-based interactions can further enhance communication between the two participants, allowing natural, intuitive commands during surgery. Here, the system can enhance precision, reduce risk, and facilitate collaboration between geographically distant medical professionals in high-stakes environments like surgeries.

CHAPTER 6

Conclusion and Future Work

6.1 Conclusion

This dissertation has explored the development of an MR-assisted, DT-enabled robot collaborative system with human-in-the-loop control. The primary objective consisted on enhancing Human-Robot Collaboration by integrating advanced technologies that bridge the gap between physical and virtual environments supporting remote collaboration and, thereby, fostering more efficient and intuitive interactions in manufacturing settings.

On-Site interaction was a crucial aspect of the project. By utilizing Handheld Devices such as tablets and smartphones, on-site participants are able to share live views of their surroundings. Furthermore, augmented reality elements were layed upon the MR application to provide visual and audio cues to the user and enhance the robot's movements awareness.

In terms of **remote visualization and interaction**, a foundational 2D interface was accessible via standard devices like laptops. This interface enabled remote participants to visualize the collaboration scenario and understand the task context effectively. Furthermore, the system's capabilities allow remote operation of the robot through the MR application, given that bi-directional communication was implemented. This enhancement empowered remote users to interact with and manipulate the collaborative environment, bringing them closer to the on-site experience, aiming to improve the overall collaborative efficacy.

Regarding **automation and immersion**, a camera was mounted on the robot to automate the process of environment sharing with remote participants. This feature relieved on-site participants from the responsibility of manually sharing visual information, as the robot could now autonomously provide live feeds of the workspace. This automation not only improved efficiency but also enhanced the immersive experience for remote users by offering real-time visual insights into the operational environment.

Throughout the development process, we leveraged the Unity game engine for robot model development and employed the Robot Operating System for seamless communication between the physical robot and its DT. The use of Vuforia facilitated precise pose registration, ensuring accurate alignment between virtual and physical models. By integrating both visual and audio cues as well as intuitive controls within the shared environment, we enhanced user awareness of the robot's movements and provided a user-friendly interface for robot manipulation.

In conclusion, the project successfully achieved its main objectives by developing a system that integrates MR and DT technologies to enhance HRC, particularly focusing on the remote participant's capabilities. While direct studies were not conducted to evaluate usability, the implemented functionalities suggest that, with further refinement and user-centered adjustments, the system holds potential to improve the intuitiveness, efficiency, and safety of collaborative environments. Future evaluations and iterative improvements will help align the system more closely with the principles of Industry 5.0, advancing human-centric and flexible industrial practices.

6.2 Future Work

Despite having achieved its primary goals, there are several areas for future exploration to further enhance the system's capabilities and impact. Future work related to this includes:

- **Enhancing Immersive Technologies**: Exploring the potential of advanced devices like the Microsoft HoloLens 2 for the on-site member and Meta Quest for the remote user can significantly enhance the immersive experience of remote collaboration. Integrating MR headsets enables spatial awareness, allowing the HoloLens 2 to map and understand the physical environment, providing more precise interactions. Combined with enhanced object manipulation and real-time data integration, users can interact intuitively within the collaborative environment, overlaying critical information and executing complex tasks with precision. For the remote user, Meta Quest provides immersive visualization, allowing them to fully engage with 3D content in a virtual space. Its mobility allows the remote user to work from different locations seamlessly. Immersive collaboration tools such as whiteboard sharing and 3D model manipulation enable richer, real-time teamwork, further boosting collaborative efficiency.
- **Improving Communication Tools**: Integrating advanced communication tools such as interactive annotations on 3D models and gesture-based interaction will further enhance collaboration between on-site and remote participants. Remote users can annotate specific areas on the digital twin in real time, providing clear, visual instructions to the on-site member. Gesture-based interaction, meanwhile, offers both users a more intuitive, natural means of communication, allowing for non-verbal cues and actions. However, this approach must be adapted to the specific requirements of each use case, as varying tasks may demand different levels of precision, feedback, or interaction.
- **Conducting User Study:** is a critical step to refine the system based on real-world usage feedback and performance metrics. By carrying out comprehensive usability studies, we can evaluate different aspects of the system separately, such as UI design, robot control modes, and task execution. This will help identify key points and areas for improvement. Furthermore, analyzing task performance data will allow us to optimize workflows, improve system ergonomics, and enhance overall efficiency, ensuring that the system is tailored to meet users' diverse needs.

- **Longitudinal Studies and Ergonomic**: Evaluating the long-term impact of the developed system on collaboration efficiency and user well-being is essential. Longitudinal studies can assess how the technology influences productivity, identifying trends that inform further improvements. Additionally, ergonomic assessments will ensure that AR devices do not cause discomfort or health issues, optimizing long-term usage for user well-being.

In addition to user experience research, future work could focus on:

- **Advanced Interaction Modalities**: Incorporating gesture recognition and voice commands can make the system more accessible and reduce reliance on manual input devices. These modalities can provide a more intuitive control mechanism, especially in environments where traditional input devices are impractical.
- **Security and Privacy Measures**: Strengthening encryption and authentication protocols will protect user data and ensure secure communication channels. This is crucial for maintaining trust and compliance with data protection regulations.

By pursuing these future developments, the system can significantly improve its effectiveness and user satisfaction. Continuous refinement based on user feedback and technological advancements could contribute to its adoption in various industrial contexts, ultimately enhancing HRC and advancing the principles of Industry 5.0.

6.3 Final Remarks

This dissertation has laid the groundwork for addressing HRC in manufacturing environments. By integrating MR and DT technologies, we have created a system that may enhance human performance and efficient, as well as enhance collaboration between users physicall distributted through the use of a new reality. The insights gained and the foundation established through this work pave the way for future explorations that can further bridge the gap between humans and machines. Embracing continuous improvement and adaptation will ensure that such systems remain relevant and impactful in the ever-evolving landscape of industrial automation.

References

[1] A. Weiss, A. K. Wortmeier, and B. Kubicek, «Cobots in Industry 4.0: A Roadmap for Future Practice Studies on Human-Robot Collaboration», *IEEE Transactions on Human-Machine Systems*, vol. 51, pp. 335–345, 4 Aug. 2021, ISSN: 21682305. DOI: 10.1109/THMS.2021.3092684.

[2] M. Speicher, B. D. Hall, and M. Nebeling, «What is Mixed Reality?», in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19, Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–15, ISBN: 9781450359702. DOI: 10.1145/3290605.3300767. [Online]. Available: https://doi.org/10.1145/3290605.3300767.

[3] D. P. Moller, H. Vakilzadian, and R. E. Haas, «From Industry 4.0 towards Industry 5.0», vol. 2022-May, IEEE Computer Society, 2022, pp. 61–68, ISBN: 9781665480093. DOI: 10.1109/eIT53891.2022.9813831.

[4] I. Ahmed, G. Jeon, and F. Piccialli, «From Artificial Intelligence to Explainable Artificial Intelligence in Industry 4.0: A Survey on What, How, and Where», *IEEE Transactions on Industrial Informatics*, vol. 18, pp. 5031–5042, 8 Aug. 2022, ISSN: 19410050. DOI: 10.1109/TII.2022.3146552.

[5] M. Golovianko, V. Terziyan, V. Branytskyi, and D. Malyk, «Industry 4.0 vs. Industry 5.0: Co-existence, Transition, or a Hybrid», *Procedia Computer Science*, vol. 217, pp. 102–113, 2023, 4th International Conference on Industry 4.0 and Smart Manufacturing, ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2022.12.206. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050922022840.

[6] 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. DOI: doi/10.2777/308407.

[7] S. Nahavandi, «Industry 5.0—A Human-Centric Solution», *Sustainability*, vol. 11, no. 16, 2019, ISSN: 2071-1050. DOI: 10.3390/su11164371. [Online]. Available: https://www.mdpi.com/2071-1050/11/16/4371.

[8] N. Y. Mulongo, «Industry 5.0 a Novel Technological Concept», in *2024 International Conference on Smart Applications, Communications and Networking (SmartNets)*, 2024, pp. 1–6. DOI: 10.1109/SmartNets61466.2024.10577684.

[9] U. E. Ogenyi, J. Liu, C. Yang, Z. Ju, and H. Liu, «Physical Human-Robot Collaboration: Robotic Systems, Learning Methods, Collaborative Strategies, Sensors, and Actuators», *IEEE Transactions on Cybernetics*, vol. 51, pp. 1888–1901, 4 Apr. 2021, ISSN: 21682275. DOI: 10.1109/TCYB.2019.2947532.

[10] R. Jahanmahin, S. Masoud, J. Rickli, and A. Djuric, *Human-robot interactions in manufacturing: A survey of human behavior modeling*, Dec. 2022. DOI: 10.1016/j.rcim.2022.102404.

[11] A. A. Malik and A. Brem, «Digital twins for collaborative robots: A case study in human-robot interaction», *Robotics and Computer-Integrated Manufacturing*, vol. 68, p. 102 092, 2021, ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2020.102092. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584520303021.

[12] J. Colgate, W. Wannasuphoprasit, and M. Peshkin, «Cobots: robots for collaboration with human operators», English (US), Proceedings of the 1996 ASME International Mechanical Engineering Congress and Exposition ; Conference date: 17-11-1996 Through 22-11-1996, Dec. 1996, pp. 433–439.

[13] E. Matheson, R. Minto, E. G. G. Zampieri, M. Faccio, and G. Rosati, «Human–Robot Collaboration in Manufacturing Applications: A Review», *Robotics*, vol. 8, no. 4, 2019, ISSN: 2218-6581. DOI: 10.3390/robotics8040100. [Online]. Available: https://www.mdpi.com/2218-6581/8/4/100.

[14] L. Barbazza, M. Faccio, F. Oscari, and G. Rosati, «Agility in Assembly Systems: A Comparison Model», *Assembly Automation*, vol. 37, no. 4, pp. 411–421, 2017. DOI: 10.1108/AA-10-2016-128. [Online]. Available: https://doi.org/10.1108/AA-10-2016-128.

[15] K. R. Guerin, C. Lea, C. Paxton, and G. D. Hager, «A framework for end-user instruction of a robot assistant for manufacturing», in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 6167–6174. DOI: 10.1109/ICRA.2015.7140065.

[16] G. Michalos, S. Makris, P. Tsarouchi, T. Guasch, D. Kontovrakis, and G. Chryssolouris, «Design Considerations for Safe Human-robot Collaborative Workplaces», *Procedia CIRP*, vol. 37, pp. 248–253, 2015, CIRPe 2015 - Understanding the life cycle implications of manufacturing, ISSN: 2212-8271. DOI: https://doi.org/10.1016/j.procir.2015.08.014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212827115008550.

[17] M. Faccio, M. Bottin, and G. Rosati, «Collaborative and Traditional Robotic Assembly: A Comparison Model», *International Journal of Advanced Manufacturing Technology*, vol. 102, pp. 1355–1372, 2019. DOI: 10.1007/s00170-018-03247-z. [Online]. Available: https://doi.org/10.1007/s00170-018-03247-z.

[18] M. Fechter, P. Foith-Förster, M. S. Pfeiffer, and T. Bauernhansl, «Axiomatic Design Approach for Human-robot Collaboration in Flexibly Linked Assembly Layouts», *Procedia CIRP*, vol. 50, pp. 629–634, 2016, 26th CIRP Design Conference, ISSN: 2212-8271. DOI: https://doi.org/10.1016/j.procir.2016.04.186. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212827116305686.

[19] P. Ladwig and C. Geiger, «A Literature Review on Collaboration in Mixed Reality», in *Smart Industry & Smart Education*, M. E. Auer and R. Langmann, Eds., Cham: Springer International Publishing, 2019, pp. 591–600, ISBN: 978-3-319-95678-7.

[20] Y. Liu, S. Ong, and A. Nee, «State-of-the-art survey on digital twin implementations», *Advances in Manufacturing*, vol. 10, pp. 1–23, 2022. DOI: 10.1007/s40436-021-00375-w.

[21] R. T. Azuma, *A Survey of Augmented Reality*, 1997. [Online]. Available: http://www.cs.unc.edu/~azumaW:.

[22] D. Chusetthagarn, V. Visoottiviseth, and J. Haga, «A Prototype of Collaborative Augment Reality Environment for HoloLens», in *2018 22nd International Computer Science and Engineering Conference (ICSEC)*, 2018, pp. 1–5. DOI: 10.1109/ICSEC.2018.8712803.

[23] Microsoft, *Mixed Reality*, https://learn.microsoft.com/en-us/windows/mixed-reality/discover/mixed-reality, Accessed: September 16, 2024, 2024.

[24] A. Mayer, J.-R. Chardonnet, P. Häfner, and J. Ovtcharova, «Collaborative Work Enabled by Immersive Environments», in *New Digital Work: Digital Sovereignty at the Workplace*, A. Shajek and E. A. Hartmann, Eds. Cham: Springer International Publishing, 2023, pp. 87–117, ISBN: 978-3-031-26490-0. DOI: 10.1007/978-3-031-26490-0_6. [Online]. Available: https://doi.org/10.1007/978-3-031-26490-0_6.

[25] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, «Digital Twin in Industry: State-of-the-Art», *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, 2019. DOI: 10.1109/TII.2018.2873186.

[26] K. H. Soon and V. H. S. Khoo, «CITYGML MODELLING FOR SINGAPORE 3D NATIONAL MAPPING», *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-4/W7, pp. 37–42, 2017. DOI: 10.5194/isprs-archives-XLII-4-W7-37-2017. [Online]. Available: https://isprs-archives.copernicus.org/articles/XLII-4-W7/37/2017/.

[27] W. Luo, T. Hu, W. Zhu, and F. Tao, «Digital twin modeling method for CNC machine tool», in *2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*, 2018, pp. 1–4. DOI: 10.1109/ICNSC.2018.8361285.

[28] F. Tao, M. Zhang, Y. Liu, and A. Nee, «Digital twin driven prognostics and health management for complex equipment», *CIRP Annals*, vol. 67, no. 1, pp. 169–172, 2018, ISSN: 0007-8506. DOI: https://doi.org/10.1016/j.cirp.2018.04.055. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0007850618300799.

[29] M. N. Mamatha, «Design of Single Patient Care Monitoring System and Robot», in *Cyber-physical Systems and Digital Twins*, M. E. Auer and K. Ram B., Eds., Cham: Springer International Publishing, 2020, pp. 203–216, ISBN: 978-3-030-23162-0.

[30] C. Doukas and I. Maglogiannis, «Bringing IoT and Cloud Computing towards Pervasive Healthcare», in *2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, 2012, pp. 922–926. DOI: 10.1109/IMIS.2012.26.

[31] F. Tao, F. Sui, A. Liu, Q. Qi, M. Zhang, B. Song, Z. Guo, S. Lu, and A. Nee, «Digital twin-driven product design framework», *International Journal of Production Research*, vol. 57, pp. 1–19, Feb. 2018. DOI: 10.1080/00207543.2018.1443229.

[32] R. Rosen, G. von Wichert, G. Lo, and K. D. Bettenhausen, «About The Importance of Autonomy and Digital Twins for the Future of Manufacturing», *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 567–572, 2015, 15th IFAC Symposium onInformation Control Problems inManufacturing, ISSN: 2405-8963. DOI: https://doi.org/10.1016/j.ifacol.2015.06.141. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896315003808.

[33] C. Cimino, E. Negri, and L. Fumagalli, «Review of digital twin applications in manufacturing», *Computers in Industry*, vol. 113, p. 103 130, 2019, ISSN: 0166-3615. DOI: https://doi.org/10.1016/j.compind.2019.103130. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166361519304385.

[34] Y. Lu, C. Liu, K. I.-K. Wang, H. Huang, and X. Xu, «Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues», *Robotics and Computer-Integrated Manufacturing*, vol. 61, p. 101 837, 2020, ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2019.101837. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584519302480.

[35] Y. K. Liu, S. K. Ong, and A. Y. C. Nee, «State-of-the-art survey on digital twin implementations», *Advanced Manufacturing*, vol. 10, no. 1, pp. 1–23, 2022. DOI: 10.1007/s40436-021-00375-w. [Online]. Available: https://doi.org/10.1007/s40436-021-00375-w.

[36] C.-H. Chu and Y.-L. Liu, «Augmented reality user interface design and experimental evaluation for human-robot collaborative assembly», *Journal of Manufacturing Systems*, vol. 68, pp. 313–324, 2023, ISSN: 0278-6125. DOI: https://doi.org/10.1016/j.jmsy.2023.04.007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278612523000717.

[37] S. A. Green, M. Billinghurst, X. Chen, and J. G. Chase, «Human-Robot Collaboration: A Literature Review and Augmented Reality Approach in Design», *International Journal of Advanced Robotic Systems*, vol. 5, no. 1, p. 1, 2008. DOI: 10.5772/5664. eprint: ://doi.org/10.5772/5664. [Online]. Available: ://doi.org/10.5772/5664.

[38] P. A. Lasota and J. A. Shah, «Analyzing the Effects of Human-Aware Motion Planning on Close-Proximity Human–Robot Collaboration», *Human Factors*, vol. 57, no. 1, pp. 21–33, 2015, PMID: 25790568. DOI: 10.1177/0018720814565188. eprint: https://doi.org/10.1177/0018720814565188. [Online]. Available: https://doi.org/10.1177/0018720814565188.

[39] S. Ong, A. Nee, A. Yew, and N. Thanigaivel, «AR-Assisted Robot Welding Programming», *Advances in Manufacturing*, vol. 8, no. 1, pp. 40–48, 2020, ISSN: 2195-3597. DOI: 10.1007/s40436-019-00283-0. [Online]. Available: https://doi.org/10.1007/s40436-019-00283-0.

[40] I. Malý, D. Sedláček, and P. Leitão, «Augmented reality experiments with industrial robot in industry 4.0 environment», in *2016 IEEE 14th International Conference on Industrial Informatics (INDIN)*, 2016, pp. 176–181. DOI: 10.1109/INDIN.2016.7819154.

[41] D. Puljiz and B. Hein, *Concepts for End-to-end Augmented Reality based Human-Robot Interaction Systems*, 2019. arXiv: 1910.04494 [cs.RO]. [Online]. Available: https://arxiv.org/abs/1910.04494.

[42] D. Puljiz, E. Stöhr, K. S. Riesterer, B. Hein, and T. Kröger, «General Hand Guidance Framework using Microsoft HoloLens», in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 5185–5190. DOI: 10.1109/IROS40897.2019.8967649.

[43] B. Tipary and G. Erdős, «Generic development methodology for flexible robotic pick-and-place workcells based on Digital Twin», *Robotics and Computer-Integrated Manufacturing*, vol. 71, p. 102 140, 2021,

ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2021.102140. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584521000259.

[44] M. Parigi Polverini, A. M. Zanchettin, and P. Rocco, «A constraint-based programming approach for robotic assembly skills implementation», *Robotics and Computer-Integrated Manufacturing*, vol. 59, pp. 69–81, 2019, ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2019.03.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584518305726.

[45] J. Jiang, Z. Huang, Z. Bi, X. Ma, and G. Yu, «State-of-the-Art control strategies for robotic PiH assembly», *Robotics and Computer-Integrated Manufacturing*, vol. 65, p. 101 894, 2020, ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2019.101894. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584519302418.

[46] G.-Z. Yang, «Robot learning—Beyond imitation», *Science Robotics*, vol. 4, no. 26, eaaw3520, 2019. DOI: 10.1126/scirobotics.aaw3520. eprint: https://www.science.org/doi/pdf/10.1126/scirobotics.aaw3520. [Online]. Available: https://www.science.org/doi/abs/10.1126/scirobotics.aaw3520.

[47] G. Boschetti, M. Bottin, M. Faccio, and R. Minto, «Multi-robot multi-operator collaborative assembly systems: a performance evaluation model», *Journal of Intelligent Manufacturing*, vol. 32, no. 5, pp. 1455–1470, 2021, ISSN: 1572-8145. DOI: 10.1007/s10845-020-01714-7. [Online]. Available: https://doi.org/10.1007/s10845-020-01714-7.

[48] F. Azimifar, M. Abrishamkar, B. Farzaneh, A. Sarhan, and H. Amini, «Improving teleoperation system performance in the presence of estimated external force», *Robotics and Computer-Integrated Manufacturing*, vol. 46, pp. 86–93, 2017, ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2016.12.004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584516301879.

[49] T. Masood and J. Egger, «Augmented reality in support of Industry 4.0—Implementation challenges and success factors», *Robotics and Computer-Integrated Manufacturing*, vol. 58, pp. 181–195, 2019, ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2019.02.003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584518304101.

[50] A. Hietanen, R. Pieters, M. Lanz, J. Latokartano, and J.-K. Kämäräinen, «AR-based interaction for human-robot collaborative manufacturing», *Robotics and Computer-Integrated Manufacturing*, vol. 63, p. 101 891, 2020, ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2019.101891. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584519307355.

[51] Y. Pan, C. Chen, D. Li, Z. Zhao, and J. Hong, «Augmented reality-based robot teleoperation system using RGB-D imaging and attitude teaching device», *Robotics and Computer-Integrated Manufacturing*, vol. 71, p. 102 167, 2021, ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2021.102167. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S073658452100051X.

[52] S. Ong, A. Yew, N. Thanigaivel, and A. Nee, «Augmented reality-assisted robot programming system for industrial applications», *Robotics and Computer-Integrated Manufacturing*, vol. 61, p. 101 820, 2020, ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2019.101820. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584519300250.

[53] M. Matulis and C. Harvey, «A robot arm digital twin utilising reinforcement learning», *Computers & Graphics*, vol. 95, pp. 106–114, 2021, ISSN: 0097-8493. DOI: https://doi.org/10.1016/j.cag.2021.01.011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S009784932100011X.

[54] E. Yigitbas, K. Karakaya, I. Jovanovikj, and G. Engels, *Enhancing Human-in-the-Loop Adaptive Systems through Digital Twins and VR Interfaces*, 2021. arXiv: 2103.10804 [cs.HC]. [Online]. Available: https://arxiv.org/abs/2103.10804.

[55] I. A. Tsokalo, D. Kuss, I. Kharabet, F. H. P. Fitzek, and M. Reisslein, «Remote Robot Control with Human-in-the-Loop over Long Distances Using Digital Twins», in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6. DOI: 10.1109/GLOBECOM38437.2019.9013428.

[56] C. Li, P. Zheng, S. Li, Y. Pang, and C. K. Lee, «AR-assisted digital twin-enabled robot collaborative manufacturing system with human-in-the-loop», *Robotics and Computer-Integrated Manufacturing*, vol. 76, p. 102 321, 2022, ISSN: 0736-5845. DOI: https://doi.org/10.1016/j.rcim.2022.102321. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584522000102.

APPENDIX A

Additional content

A.1 Unity Robotics Hub Overview

Before diving into the specifics of establishing the ROS-Unity connection and further develop the project, both tutorials and resources available in the Unity Robotics Hub were studied. This GitHub repository serves as a central hub for tools, tutorials, and documentation tailored for robotic simulation in Unity.

A.1.1 Available Documentation

It offers a range of tutorials that are invaluable for setting up and extending ROS-Unity integration, as well as to understand how ROS concepts work inside Unity's environment:

- **ROS–Unity Integration: Initial Setup** - Guides you through the initial steps of setting up communication between ROS and Unity, including package installation and network configuration.
- **ROS–Unity Integration: Network Description** - Provides a detailed overview of network settings and offers troubleshooting tips for connectivity issues.
- **ROS–Unity Integration: Publisher** - Teaches you how to publish messages from a Unity scene to ROS, with practical examples involving GameObject data.
- **ROS–Unity Integration: Subscriber** - Demonstrates how to subscribe to ROS topics in Unity and use the received messages to alter objects in a Unity scene.
- **ROS–Unity Integration: Unity Service** - Covers the implementation of ROS services within Unity, allowing Unity to respond to ROS service requests.
- **ROS–Unity Integration: Service Call** - Explains how to call external ROS services from Unity, enabling Unity to request data or actions from ROS nodes.

The repository also includes example scripts that correspond to each tutorial.

A.2 Establishing the Network Connection

After reviewing the Unity Robotics Hub tutorial on network integration, it became clear that establishing a network connection between the Unity and ROS environments was the first crucial step in remote application development. The process involves:

- **Setting up the network:** Connect the Unity laptop to a Wi-Fi network, then connect the Ubuntu laptop running ROS to the hotspot created by the Unity laptop.
- **Configuring the IP address:** Use the IP address from the Unity laptop within the Unity inspector as shown in Figure A.1, and in the `ROSConnection.cs` script to ensure proper communication.
- **Specify the IP Address in ROS Workspace:** A new *.launch* file was created to initialize new nodes, including the `server_endpoint` node from the `ros_tcp_endpoint` package, crucial for establishing a proper connection between the ROS and Unity environments. An example of the IP definition in the *.launch* file can be seen below.

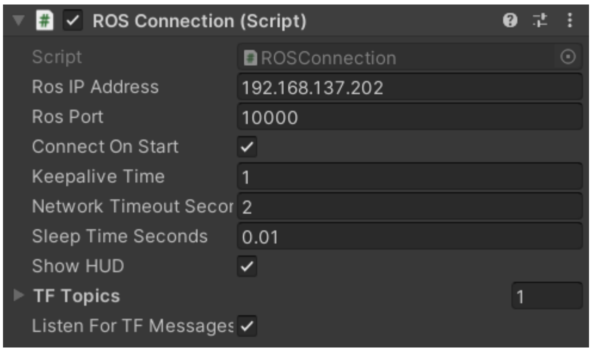


Figure A.1: Unity Connection Inspector

```
<arg name="tcp_ip" default="192.168.137.202"/>
<arg name="tcp_port" default="10000"/>

<node name="server_endpoint" pkg="ros_tcp_endpoint"
    type="default_server_endpoint.py" args="--wait" output="screen"
    respawn="true">
    <param name="tcp_ip" type="string" value="$(arg tcp_ip)"/>
    <param name="tcp_port" type="int" value="$(arg tcp_port)"/>
</node>
```

This setup is fundamental for the Unity environment to interact effectively with ROS, allowing for real-time data exchange and control commands to be sent between the two systems. Further details are available in the Unity Robotics Hub tutorial on ROS-Unity integration.

A.3 ROS Message Generation

After establishing the communication between ROS and Unity, as well as the other basic communication channels—publishers, subscribers, and services, the next step was to customize these components to handle the specific types of messages required for this project.

A.3.1 Adapting to Specific Message Types

To begin, I revisited the previously mentioned GitHub repositories for the UR10e robot, specifically focusing on identifying which ROS nodes and topics were critical for my Unity-ROS integration. This involved:

- **Identifying Key Nodes and Topics:** The primary focus was on finding which topic is responsible for publishing the current state of the robot's joints.
- **Message Subscription and Retrieval:** With guidance from my supervisors, I determined that subscribing to the `/joint_states` topic would be essential for retrieving real-time data about the robot's joint positions. This topic uses the `sensor_msgs/JointState` message type, a standard message in ROS that provides the positions, velocities, and efforts of the robot's joints. It is well-documented and includes the necessary fields for capturing joint states. This message type is part of the broader `common_msgs/sensor_msgs` package, which is available on the ROS wiki and in the sensor_msgs GitHub repository.

A.3.2 ROS-TCP Connector and C# Message Generation

To correctly handle the `sensor_msgs/JointState` messages within Unity, it was necessary to generate corresponding C# classes. This process, which is detailed in the ROS-TCP Connector documentation, involves:

1. **Generating C# Message Classes:** By following the steps provided in the documentation, I was able to generate the necessary C# JointState class that represent the corresponding ROS message, as depicted in the figure A.2. This step was crucial for storing and manipulating the joint state data within the Unity environment.

Figure A.2: JointState message generation in Unity, corresponding to the desired ROS message

2. **Compiling and Verifying the Message Classes:** After generating the message classes, I compiled them in the Windows environment, ensuring they matched the expected structure as described in the UnityRoboticsTutorial repository. This process took some time to fully understand, but was critical for the successful integration of ROS data into Unity.

A.4 Position Control (1 control type) - Unity to ROS

- confirm the purpose of this section - remove it? This control type enables operators to interact directly with the robot's joints in real-time, providing a responsive and highly interactive control environment. The implementation of Position Control involves several key steps to ensure seamless operation and integration with the ROS environment.

A.4.1 Saving and Sending Joint States

The first step in implementing Position Control was to capture and save the current states of the robot's joints in Unity. These joint states are then packaged and sent to the ROS environment upon user command. This process involves:

1. **Capturing Joint States:** The Unity application continuously monitors and records the positions of each joint of the digital twin robot, visible in the debug console represented in the figure A.3

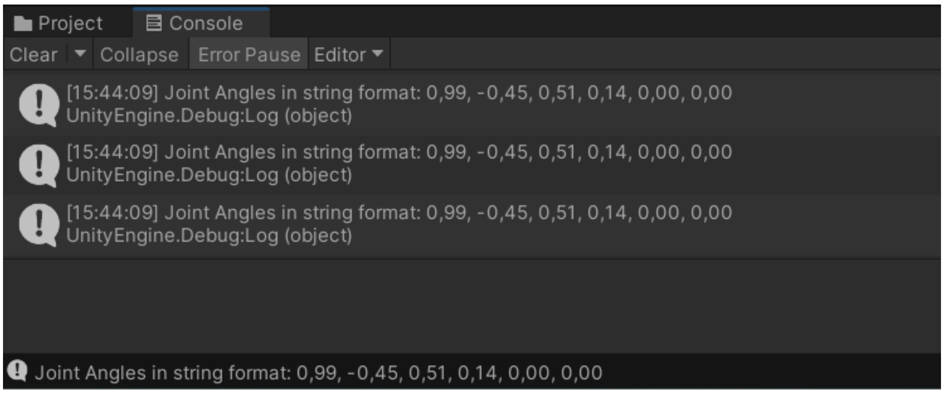


Figure A.3: Unity debug console showing the current digital twin's joint positions

2. **Data Serialization:** The joint states are serialized into a format suitable for transmission over the network. This was done by updating the joint positions into a list that is constantly saved into a JSON format, as depicted in the picture A.4.
3. **Sending Data:** Upon pressing a UI button trigger the serialized joint states are sent from Unity to the ROS environment using the previously established TCP/IP connection.

```json
{
    "jointPositions": [
        {
            "jointName": "shoulder_pan_joint",
            "position": 0.9948359727859497
        },
        {
            "jointName": "shoulder_lift_joint",
            "position": -0.4398227035999298
        },
        {
            "jointName": "elbow_joint",
            "position": 0.5078903436660767
        },
        {
            "jointName": "wrist_1_joint",
            "position": 0.1413716971874237
        },
        {
            "jointName": "wrist_2_joint",
            "position": 0.0
        },
        {
            "jointName": "wrist_3_joint",
            "position": 0.0
        }
    ]
}
```


Figure A.4: JSON format used to store Unity's digital twin joint states

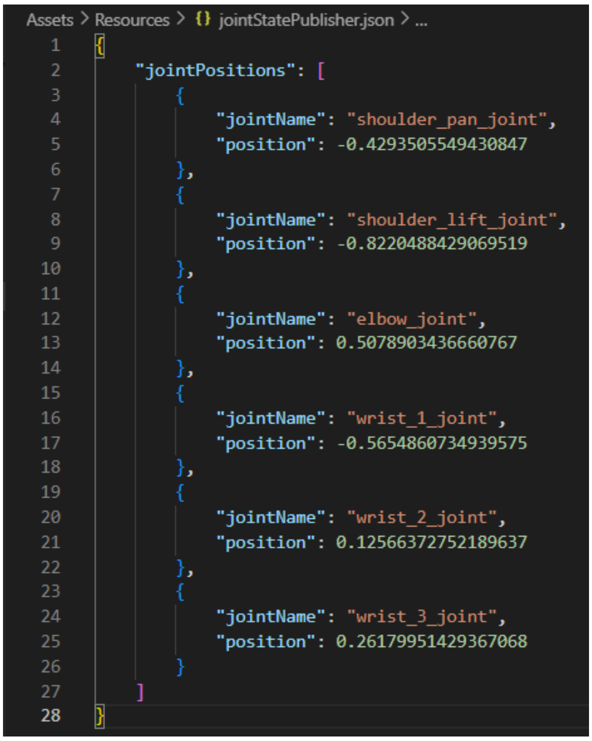
A.4.2 ROS Integration for Position Control

In order to receive the data, the ROS environment needed some adjustments to process these joint states coming from Unity environment. Two new nodes were added to the ROS setup, enabling the following key functionalities:

1. **Joint State Listener Node:**
   - **Purpose:** Receives joint states from Unity, ensuring that the digital inputs are translated into actionable data within the ROS environment.
   - **Functionality:** This node listens to a dedicated topic from Unity, processes this data, and republishes it to a different control topic. Figure A.5 illustrates an example of data transmitted from Unity to ROS. The accompanying debug log below demonstrates its accuracy.

```
rosrun iris_ur10e unity_joint_subscriber.py
[INFO] [1725629583.306162, 720.978000]:
/joint_state_listener_7441_1725629572125I
heard (-0.4293505549430847, -0.8220488429069519,
0.5078903436660767, -0.5654860734939575,
0.12566372752189636, 0.26179951429367065)
```

2. **Robot Control Node:**
   - **Purpose:** Directly controls the robot's movements based on processed joint states from Unity.



Assets > Resources > {} jointStatePublisher.json > ...

```json
{
    "jointPositions": [
        {
            "jointName": "shoulder_pan_joint",
            "position": -0.4293505549430847
        },
        {
            "jointName": "shoulder_lift_joint",
            "position": -0.8220488429069519
        },
        {
            "jointName": "elbow_joint",
            "position": 0.5078903436660767
        },
        {
            "jointName": "wrist_1_joint",
            "position": -0.5654860734939575
        },
        {
            "jointName": "wrist_2_joint",
            "position": 0.12566372752189637
        },
        {
            "jointName": "wrist_3_joint",
            "position": 0.26179951429367068
        }
    ]
}
```

Figure A.5: JSON file containing the unity's joint states that were sent to ROS environment

- **Functionality:** This node requires that the previous one is also running, and subscribes to the control topic */move_joint_unity* to fetch and apply joint state data to the physical robot, mirroring the Unity operator's interactions. The following debug log outputs the correct utilization of this node

```
rosrun iris_sami move_unity.py
[ INFO] [1725629576.546376282]: Loading robot model 'ur10e'.
[ INFO] [1725629577.773871662, 715.485000000]: Ready to take
commands for planning group manipulator.
[INFO] [1725629583.310087, 720.982000]: Moving arm to joint
positions: (-0.4293505549430847, -0.8220488429069519,
0.5078903436660767, -0.5654860734939575,
0.12566372752189636, 0.26179951429367065)
```

and the figure A.6 displays two cases where the robot and its digital twin are properly synchronized using this control type.

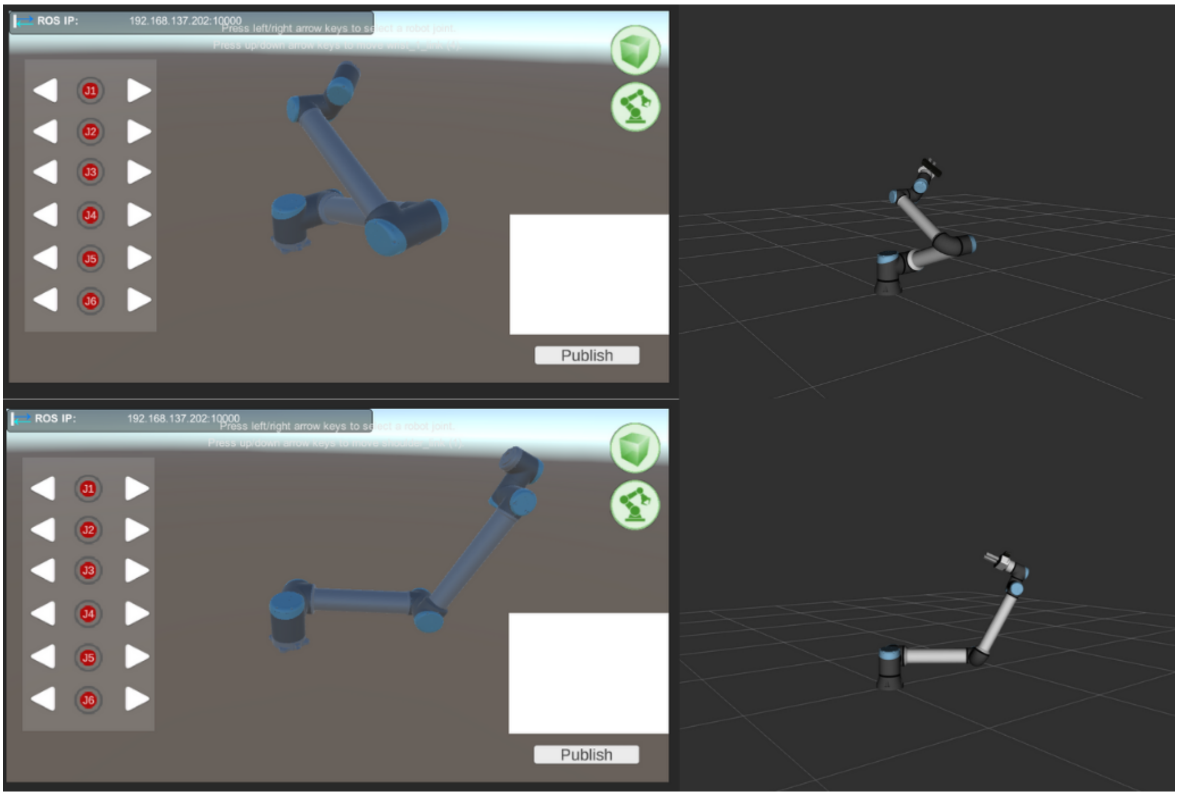


Figure A.6: Two scenarios showcasing the synchronization between Unity and ROS environments using Joint Control

A.5 Joint State Subscription (2 control type) - ROS to Unity

While the Position Control type allows operators to interactively manipulate the robot's joints and send these changes to the ROS environment, the Joint State Subscription control type operates in an opposite manner. It is designed to continually update the Unity digital twin's joint positions in synchronization with movements from the physical robot or its simulation in RViz.

A.5.1 Saving ROS Data

In order to properly synchronize Unity's digital twin with real-time robot movements from the ROS environment, a new script `JointStateSubscriber.cs` was created. It subscribes to the */joint_states* topic to continuously capture and store the robot's joint positions. This process is critical for maintaining a live reflection of the robot's state within the Unity simulation.

Subscribing to ROS Topics and Data Serialization

- **Subscribing to Joint States:** The script actively listens to the `/joint_states` topic, ensuring that any movement in the robot is promptly reflected in Unity.
- **Storing Joint Data:** Captured joint positions are stored in a C# dictionary, facilitating efficient data access and manipulation.
- **Serializing Data to JSON:** The joint data is serialized into a JSON file, which provides a persistent and accessible format for storing the robot's state, overcoming potential restrictions from Unity packages that limit direct folder access.

Explanation and Integration

The implementation of this system ensures that Unity's digital twin is consistently updated with the latest joint states from ROS, offering an accurate virtual representation for monitoring or interaction. Using the `SaveJointPositionsToFile()` method, joint data is structured and saved to enhance development workflow and project scalability.

Practical Application and Visualization

The real-time state of the robot's joints, whether it is operating in a simulated environment or in real-time with the physical robot, is represented in figure A.7. Initially, the script was designed to save these robot joints' values upon pressing a UI button, but it was later adapted to continuously update the `float64[] position` array from the ROS side, visible in the below terminal log snippet, into the Unity environment as shown in the figure A.8, ensuring that the most recent joint positions were always available.

```
name:
    - elbow_joint
    - left_finger_joint
    - right_finger_joint
```

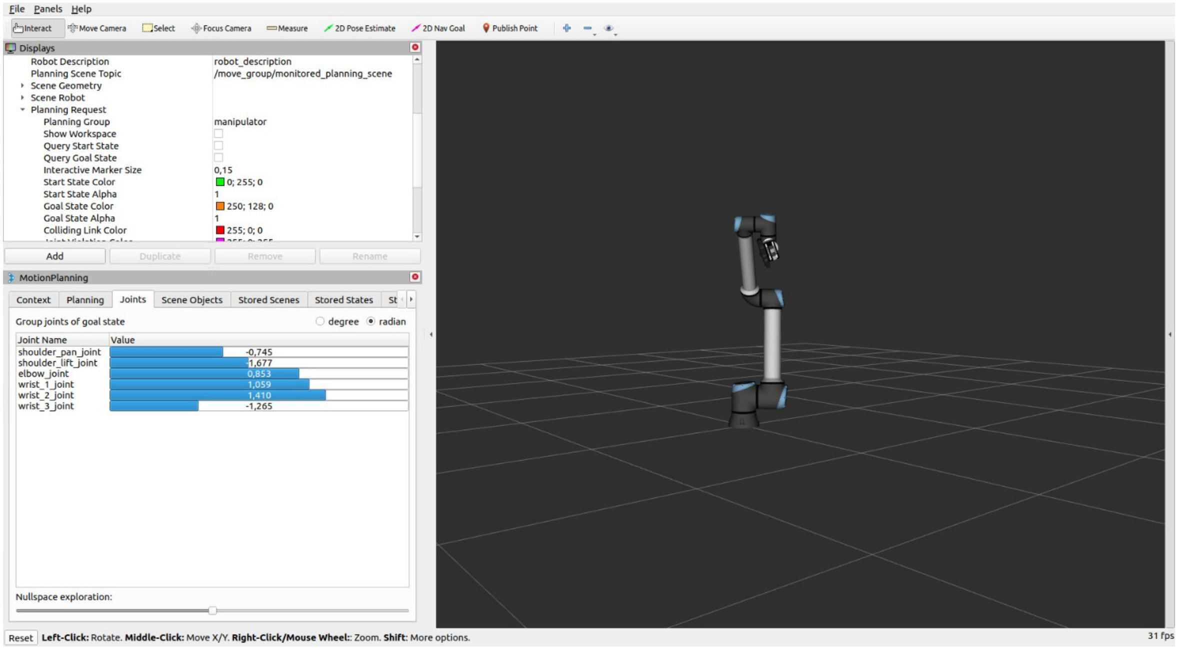
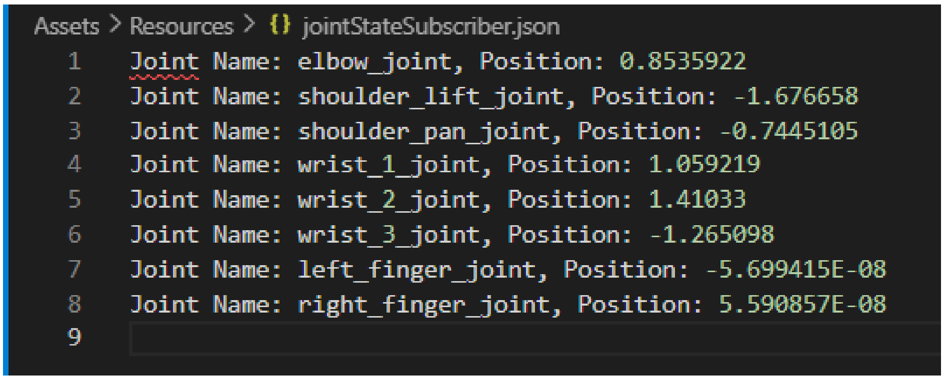


Figure A.7: Rviz environment with simulated Robot and correspondent joint values, in radians

```
    - shoulder_lift_joint
    - shoulder_pan_joint
    - wrist_1_joint
    - wrist_2_joint
    - wrist_3_joint
position: [0.8535921338670764, -5.695760001087214e-08, 5.587315793023694e-08,
-1.6766575764811593, -0.7445104810535929, 1.059219605892937, 1.410329834615414,
-1.265097896821196]
velocity: [-4.84869176906394e-05, -0.00020121543757810655, 0.0001987088475458378,
0.006705746353461187, 0.00061522726960112, -0.0003380421322876756,
0.001088906936911304, 0.0002984877856849871]
effort: [0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]
```



```
Assets > Resources > {} jointStateSubscriber.json
1  Joint Name: elbow_joint, Position: 0.8535922
2  Joint Name: shoulder_lift_joint, Position: -1.676658
3  Joint Name: shoulder_pan_joint, Position: -0.7445105
4  Joint Name: wrist_1_joint, Position: 1.059219
5  Joint Name: wrist_2_joint, Position: 1.41033
6  Joint Name: wrist_3_joint, Position: -1.265098
7  Joint Name: left_finger_joint, Position: -5.699415E-08
8  Joint Name: right_finger_joint, Position: 5.590857E-08
9
```

Figure A.8: JointStateSubscriber Json file with Robot's joint values, in radians