humantech

D5.2 – Scientific report on hybrid haptic-visual feedback mechanisms for efficient teleoperation of demolition robots

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement n° 101058236. This document reflects only the author's view, and the EU Commission is not responsible for any use that may be made of the information it contains.

Project Title	Human-Centred Technologies for a Safer and Greener European Construction Industry.
Project Acronym	HumanTech
Grant Agreement No	101058236
Instrument	Research & Innovation Action
Topic	HORIZON-CL4-2021-TWIN-TRANSITION-01-12
Start Date of Project	June 1, 2022
Duration of Project	36 months

Name of the Deliverable	Scientific report on hybrid haptic-visual feedback mechanisms for efficient teleoperation of demolition robots
Number of the Deliverable	5.2
Related WP Number and Name	WP5 – Construction Robotics and Human-Robot Collaboration
Related Task Number and Name	T5.2 - Remote interfaces for demolition
Deliverable Dissemination Level	Public
Deliverable Due Date	November 2023
Deliverable Submission Date	November 2023
Task Leader/Main Author	Alfonso Dominguez (TECNALIA)
Contributing Partners	TECNALIA, CATENDA, RPTU, SINTEF, STAMTECH, RICOH
Reviewer(s)	Jason Raphael Rambach (DFKI)

Keywords

Bilateral teleoperation, Visual feedback, Force feedback, Demolition, Haptic interface, Excavator, Admittance controller

Revisions

Version	Submission date	Comments	Author
v1.2	10-10-2023	First version	Alfonso Dominguez (TECNALIA)

Disclaimer

Acronyms and definitions

Acronym	Meaning
DoF	Degree of Freedom
BIM	Building Information Model
HW	Hardware
SW	Software
TCP	Tool Centre Point
IFC	Industry Foundation Classes
BCF	BIM Collaboration Format
HMI	Human Machine Interface

Abstract

The aim of task 5.2 “Remote interfaces for demolition is the development of a user-centric control console for the remote operation of robots in the hazardous context of demolition. This console will integrate advanced features in viewing and haptics which will increase the sense of telepresence while augmenting the precision and dexterity of the operator.

This deliverable contains a review of studies and techniques oriented to improve the performance and safety in teleoperation tasks with special focus on visual and force feedback. Some of them have been integrated in the teleoperation framework developed in HumanTech which is described in the las point of this document.

The HumanTech project

The European construction industry faces three major challenges: increase the safety and wellbeing of its workforce, improve its productivity, and become greener, making efficient use of resources.

To address these challenges, HumanTech proposes to develop **human-centred cutting-edge technologies** such as wearables for workers' safety and support and robots that can harmoniously coexist with human workers while contributing to the ecological transition of the sector.

HumanTech aims to achieve major advances in cutting-edge technologies that will enable a safe, rewarding and digital work environment for a new generation of highly skilled construction workers and engineers.

These advances will include:

- **Robotic devices equipped with vision and intelligence** that allow them to navigate autonomously and safely in highly unstructured environments, collaborate with humans and dynamically update a semantic digital twin of the construction site in which they are.
- **Smart, unobtrusive workers protection and support equipment**. From exoskeletons activated by body sensors for posture and strain to wearable cameras and XR glasses that provide real-time workers' location and guidance for them to perform their tasks efficiently and accurately.
- An entirely new breed of **Dynamic Semantic Digital Twins (DSDTs) of construction sites** that simulate in detail the current state of a construction site at the geometric and semantic level, based on an extended Building Information Modelling (BIM) formulation that contains all relevant structural and semantic dimensions (BIMxD). BIMxDs will act as a common reference for all human workers, engineers, and autonomous machines.

The **HumanTech consortium** is formed by 22 organisations — leading research institutes and universities, innovative hi-tech SMEs, and large enterprises, construction groups and a construction SME representative — from 10 countries, bringing expertise in 11 different disciplines. The consortium is led by the German Research Center for Artificial Intelligence’s Augmented Vision department.

Contents

Figures

1. INTRODUCTION

The European construction industry faces three major challenges: improve its productivity, increase the safety and wellbeing of its workforce and make a shift towards a green, resource efficient industry. To address these challenges adequately, HumanTech proposes a human-centered approach, involving breakthrough technologies such as wearables for worker safety and support, and intelligent robotic technology that can harmoniously co-exist with human workers while also contributing to the green transition of the industry.

One of HumanTech's scientific-technological objectives is converting demolition tasks in "remotely supervised semi-automated" processes to improve their productivity and security. A demolition robot will autonomously carry out tasks with the intervention of a human operator when it is needed through the teleoperation. The task 5.2 "Remote interfaces for demolition" will develop user-centric control console for remote operation of robots in the hazardous context of demolition. Advanced capabilities in viewing will be combined with novel haptic control interfaces to provide an increase sense of telepresence and situational awareness while augmenting the precision and dexterity of the remote operator. In this sense, this document will explore the state-of-the-art techniques and technologies applied in hybrid consoles and teleoperation systems to check which ones are worth including in HumanTech system.

A teleoperation system consists of at least one master robot locally manipulated by an operator, and at least one slave robot that remotely mimics the manoeuvres of the master robot to perform the operation on an environment. A communication network connects the master and the slave robots, transferring necessary information between the two sites [1].

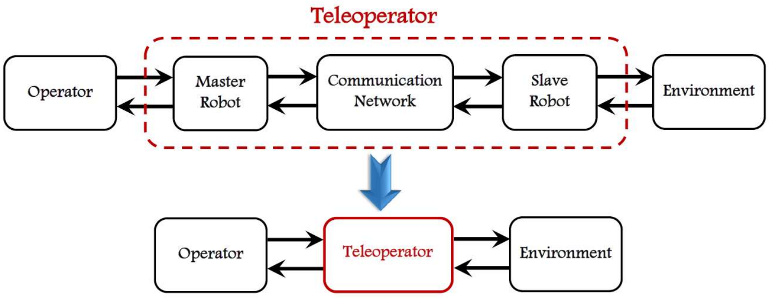


Figure 1 – Components of a single-master/single-slave teleoperation system [1]

The main objective of a teleoperation system is to extend the human capability of manipulating objects remotely by providing the operator with similar conditions as those in the remote location. In a general setting, the human imposes a force on the master manipulator which in turn results in a displacement that is transmitted to the slave that mimics that movement. If the slave possesses force sensors, then it can transmit to the master reaction forces from the task being performed, which enters the input torque of the master, and the teleoperator is said to be controlled bilaterally [2]. Although transmitting force feedback enables the user to rely on his/her tactile sense together with visual sense, this could lead to instabilities if delays are present in the communication channel. This problem is still today one of the main challenges faced by researchers in teleoperation systems.

Teleoperated robots are used in a wide variety of fields. They are mostly used in scenarios where the environment is too dangerous for humans or simply inaccessible (bomb disposal, handling of radioactive samples in laboratories or nuclear plants, disaster recovery operation, space exploration). They are also used in robotic surgery. Here the advantages are multiple: higher precision, ability to access smaller places, perform more refined motions than a human (tremor suppression) and accomplish the surgery in a more ergonomic posture.

In the last decade, several companies introduced teleoperated demolition machines. These machines rely on the presence of an operator nearby. New interaction models need to be proposed to increase the safety and productivity of these robots.

This document explores the integration of haptic-visual feedback mechanisms in the teleoperation of demolition robots, aiming to address the challenges faced by operators and enhance overall system performance. By delving into the technical aspects, design considerations, and potential applications of such feedback systems, HumanTech aims to contribute to the ongoing discourse on advancing teleoperation capabilities in the field of robotics. The document is organized as follows:

- Chapter 1 – Introduction: This chapter will introduce the HumanTech project, the problem to be tackled and the structure of the document.
- Chapter 2 - Human perception issues: This point will explore the limitations and challenges related to human perception in teleoperation scenarios and discuss how these perception challenges can impact the effectiveness and safety of teleoperating demolition robots.

- Chapter 3 - User interface solutions: This chapter will present various user interface solutions designed to address human perception issues in the context of demolition robot teleoperation focusing on visual and haptic channels.
- Explore advancements in interface design, graphical representations, and other visual aids to enhance the operator's understanding of the robot's surroundings.
- Chapter 4 - Haptic devices: This point will focus on the element for providing tactile feedback to operators during teleoperation. It will discuss the types of haptic devices suitable for demolition robot teleoperation, their technical specifications and how to overcome their limitations.
- Chapter 5 - Demolition robots: This chapter will provide an overview of the current state-of-the-art in demolition robot technology, emphasizing their capabilities, applications, and challenges.
- Chapter 6 - Teleoperation framework in HumanTech: This point will introduce the teleoperation framework developed in HumanTech to address the interaction between humans and demolition robots.
- Chapter 7 – Conclusions and future steps: This chapter will summarize the key findings and insights from the document. It will also present how the results of this document will be exploited.

2. HUMAN PERCEPTION ISSUES

There are many issues related to the human´s performance that can be encountered during teleoperation due to the lack of telepresence. So that, it can be said that teleoperation is challenging, as far as the operator´s performance in teleoperation is limited by the operator´s motor skills, ability to maintain the awareness of the situation or the difficulty to build mental models of the remote environment. In other words, most of the times, in teleoperating environments, the natural perceptual processing is not coupled adequately from the physical environment, and as a result the human perception tends to be inadequate. By this way, the teleoperator´s perception affordance in the remote scene is altered and problems in remote perception, known as scale ambiguities, arise [3].

Human performance issues that are involved in teleoperation fall in the categories of remote perception and remote manipulation in which both navigation and manipulation tasks are included. In addition, now, some of the main factors that affect the remote perception and manipulation in teleoperation and telesurgery are described, as well as some of the innovative techniques and technologies that have been designed with the aim to increase the telepresence and enhancement of the teleoperator´s performance [3].

2.1. Factors that affect remote perception and manipulation in teleoperation

The field of view (FOV), the orientation the teleoperator senses, the camera viewpoint, the depth perception, the quality of the video display and the time in which the information is fed back to the teleoperator are crucial factors that have impact on teleoperation systems and coming up each factor is explained [3].

- **Limited FOV:** Sometimes it is possible to encounter the “keyhole” effect while making use of the cameras that display the environment where the robot is navigating. Better said, that the teleoperation system´s remote environment cannot be displayed totally to the operator, and consequently, an extra effort such as the manipulation of cameras is required to solve this problem and gain a feasible situation awareness in comparison to the direct viewing.
- **Orientation:** It is very important for the teleoperator to have a good sense of orientation, because otherwise it would be impossible to navigate successfully in the remote environment. The orientation must be adequate both globally and

locally. Globally, the areas of interest related to the robot´s location must be known by the teleoperator. Locally, the teleoperator should negotiate local turns and avoid obstacles so that it is possible for him or her to navigate to the desired robot´s destinations. As well, it is indispensable to be conscious of the attitude that the robot has. It seems that with a gravity referenced view (GRV) display it is possible to the teleoperator to be more situationally aware of the robot´s attitude. For that, the alignment between the visual-hand frames must be ensured.

- **Camera viewpoint:** For robots which have end effectors, cameras can be placed on the manipulator´s gripper or end-effector and by this way, provide an egocentric view of the remote scene. Otherwise, cameras can also be placed on the robot´s body and therefore, capture the exocentric view of the manipulator´s movement. Thus, depending on where the cameras are placed, and because of the viewing angles that do not result natural for the human, the remote perception could be deteriorated.
- **Degraded depth perception:** Operators underestimate distances to objects, because of the lack of cues available in the teleoperation environment. The main problem is that there is no additional information of structure that allows for more precise depth perception through the display. Thus, the problem of spatial perception in teleoperation arise because of the confined skill that the operator has to appropriately scale objects that are known as scale ambiguities, as it is aforementioned. Scale ambiguities occur because the video feed is a 2D representation of a 3D space. Thus, the depth of the environment is compressed, and relevant information about depth and size specifications is missed or lost. That is, several problems are presented due its nonlinear nature and the difficulty in deriving 3D information [4]. In addition, depth perception problems ascend when projecting a 3D depth information onto a 2D display and consequently, the depth perception that is degraded has a negative impact on the teleoperator´s performance; the mission effectiveness is deteriorated.
- **Degraded video image:** The effective perception of the remote environment is favoured by the adequate interaction channel between the teleoperator and the robot. Additionally, factors like distance, obstacles or electronic jamming might cause challenges for the signal's strength maintenance. How the variance of teleoperator´s task performance is with changes in frame rate, resolution and grey scale must be considered. The product of these three parameters determines

the bit rate of the image displayed, and this has a direct correlation with the teleoperator´s performance. It seems that the teleoperative task performance is a bit enhanced if the frame rate exceeds 15 Hz and results impossible in the case it is taken under 5.6 Hz [5]. As a result, higher the quality of video feeds, higher the attitude of confidence or security with which the teleoperator actuates will be while teleoperating. By this way, the necessary cues to build teleoperator´s mental models (distance and size estimation) of the remote environment would not be left out. As well, a reduced bandwidth of the image display hinders spatial orientation´s maintenance.

- **Time delay:** Latency is the delay between input action and output response. This is occasioned due to the transmission across a communications network. When the latency of a system exceeds one second, the teleoperators does not continue commanding or trying to compensate for the delay anymore. Instead, he or she switches the strategy of control to a “move and wait” one. So, the latency has a negative effect on the task that the teleoperator performs.

3. USER INTERFACE SOLUTIONS

Many sensorial channels such as the visual, auditory or haptics are the components of interaction of HMI in teleoperation systems. Visual, auditory, and haptic modalities are common in those systems. On the contrary, the smell and taste feedback are still at a very early stage of research [4]. All those sensorial channels contribute in the increasement of telepresence in teleoperation system.

Here, in order to increase the enhancement of the teleoperator´s performance, some potential user interface solutions would be surveyed, that are known as multimodal displays.

The multimodal displays are those displays that have the capacity of feeding back two different types of sensory information to the teleoperator, so that he or she receives the visual data of the remote environment combined with audio or haptic data [3]. When an operator embodies a remote robot/machine using a monitor, he/she only experiences the environment visually. By using other modalities, the task performance can be improved because his/her visual cortex workload is reduced [5]. However, this requires the synchronization between them, otherwise, the feeling of immersion is reduced [6]. Immersion can be also negatively influenced by the so-called “Uncanny valley” – a break in immersion when an artificial being appears too realistic, causing negative responses toward it [7]. Focusing on haptics, a similar effect is caused when the haptic feedback does not coincide with the other modalities reducing the perception of realism [8].

Traditionally, the communication has been through graphical user interfaces in combination with multiple cameras. However, this is not satisfactory for more complex system due to the huge amount of information to be communicated to the remote operator. Thus, a better solution is to utilize multiple human senses to communicate the information; thus, use hearing, smelling, tasting, touching and vision.

However, an interesting solution, is to apply sensor bridging; thus, bridging visual information through e.g. sound. This is commonly used by parking assistants for modern cars, where actually the sound, is communicating distance and position of obstacle. By manipulating the sound, the driver can easily know possible collisions. Similar functionality can also be used in remote operation of industrial robot systems. In another words; information can be transferred across the common sensor modalities.

Typically, touch displays are used as human-robot interfaces in industrial applications. However, research focuses on different interface technologies, such as gestures, speech or multimodal approaches to create more intuitive interaction [9]. Speech control is mainly used as an interface for controlling robot operation commands, an example can be controlling robot operation commands in an assembly task. Non-verbal communication is especially suitable for noisy environments, such as construction sites. Gesture interfaces were also observed to be most popular among workers [9].

Gesture control of robots is used in robot guidance tasks, where gestures are used to define relative or absolute approximate poses of the robot. A gesture control application can either be offline or online, in offline applications a human can pre-program a robot trajectory based in gestures and execute the program later. In online applications, a human can control the robot in real-time based on gestures [10]. Gesture-based remote human-robot interaction can, for example, be achieved by using a Kinect sensor for teleoperation the robotic arm [11], or thermal-image based hand gesture recognition for human-robot collaboration on construction sites [12]. Other examples have been presented in [9].

This point will focus on audiovisuohaptic interfaces as these human senses contribute the most towards embodiment [3] [6] [13].

- **Visual and Audio Displays:** Audio cues turn to be beneficial supplements to visual feedback. This is because when the visual channel is heavily loaded, they have the capacity to make an enhancement of the awareness in the surroundings, cue visual attention and convey a variety of complex information. As a result, audio cues in combination with visual data are powerful tools in teleoperation as they facilitate teleoperators making decisions.
- **Visual and Haptic Displays:** Tactile and force displays are haptic displays that make use of pressure or vibration simulators which interact with the skin. These types of displays can provide warnings and communicate information related to the orientation, direction and teleoperator´s position and velocity.

It is true that the components and methods to visualize data, and above all visual data, can improve task performances in teleoperation as the stereo vision leads to the improvement of some depth cues´ perception, and consequently the telepresence turns out to be enhanced [14].

So even that the image and/or video display (visual data) seems to be the most important data form for the operator (it contributes to around 70% of overall human perception [13]), there are other sensory channels that could also be relevant if included. Thus, apart from the visual data, the implementation of systems such as the tactile and force feedback must also receive attention in teleoperation systems. In addition, both can offer to the operator quantitative, qualitative, or holistic knowledge about the data located in the remote zone, leading to a better and more efficient task performance [14].

Between auditory sense and haptics, the latest seem to have more impact on manipulation tasks. In [15], the performance was increased by a further 10% when visual feedback was combined with haptics. Auditory sense, on the contrary, had no significant effect on any measure when added to visual feedback. Consequently, for the performance of this project, the use of visual and haptic displays will be taken into consideration as far as the multimodal displays described before are concerned and leaving apart the audio displays.

Putting the focus on the haptic feedback, its meaning is complemented by the force, tactile and proprioceptive feedback. The first one is used to obtain the data on a virtual object hardness, weight and inertia. The second one provides the user with the virtual object´s surface contact geometry, smoothness, slippage, and temperature feel. The third one, informs about the data that is concerned to the user´s position or posture [4].

Furthermore, haptic feedback turns to be a vital sensorial modality and therefore, to provide such sensorial data in teleoperation systems, haptic interfaces are needed; desk-top or portable special-purpose hardware. Thus, haptic feedback interfaces comprise both, force and tactile feedback devices. Besides that, in order to modulate physical interactions, precise collision detection, real-time force computation, and high control-loop bandwidth must be taken into consideration [4].

3.1. Visual and haptic displays in Teleoperation

The main problem that comes up on teleoperation systems, as it is aforementioned, is that it is difficult to simulate the awareness of the remote environment to the teleoperator. This means that it is vital to define which type of sensory feedback is enough to simulate to the operator, to provide him or her an advanced telepresence [16].

Thus, in order to be possible to the teleoperator to perform the task as he or she was present in the remote zone, it has been analysed that visual and haptic sensory feedback are vital requirements; the combination of both will lead the teleoperator to perform the task as fluently, naturally and comfortably possible [16].

The information flow from the HMI to the operator is understood as the most relevant sensor transformation within a master-slave teleoperation system. Thus, the visual force-reflected teleoperation is transformed via two different ways. On the one hand, the visual modality detects the retina´s photons and on the other hand, the haptic modality tests changes in pressure on the skin and the joint, muscle, and tendon receptors of the teleoperator. Both modalities lead to recover the properties of the objects and by the use of them the central nervous system fuses such information [16].

Thanks to cues like the binocular disparity, shape-from-shading, knowledge of everyday objects with which to establish scale, texture gradients and motion parallax, the brain is capable to recreate 3D visual information. All those information sources are combined by the visual system, so that a useful 3D structure from everyday scenes is created. Therefore, teleoperation systems facilitate to the teleoperator suboptimal visual information and as a result, the performance of the teleoperator varies under distinct parameters related to the visual feedback [16].

But apart from the visual cues, the haptic ones are also helpful requirements as they influence the accuracy of the 3D information that the teleoperator recreates. The force feedback is commonly used to infer when objects are in contact and tactile feedback to infer the surfaces´ texture. All this information is conveyed by the pressure receptors, named Pacinian corpuscles, that are located under the skin and concentrated at the fingers´ tips [16].

The visual and haptic sensory feedback information is not only useful for the adjustment of movements that fail to attain their targets, but also for the correction of the stored information related to the environment and/or effector system that is not accurate. A teleoperator´s task is affected under different force and visual feedback regimes. In addition, once the slave arm that is being controlled by the teleoperator´s position commands gets in contact with an object, the forces and torques that act on the gripper or end-effector will be detected by a force/torque sensor. Hence, the torque/force readings are fed back to the master input device [16].

3.2. Sensory substitution, visual force feedback

A schematic representation of a typical teleoperation system is shown in Fig.2 and consists of a teleoperator, which includes the master robot, communication network and a slave robot, with signals transmitted between the operator and environment ends. Depending on the communication between the operator and the environment, the structure can be classified as unilateral, bilateral, or multilateral architectures. This point focuses on bilateral teleoperation systems, more concretely on those ones able to transmit the force applied in the environment to the master.

If in a bilateral teleoperation system the slave possesses force sensors, then the forces from the task being performed can be reflected to the master and displayed through a haptic interface (haptic bilateral teleoperation system). Using the haptic feedback, the operator seals the control loop by transmitting control signals and perceiving/monitoring the remote objects (slave) mediated by haptic sensation [17].

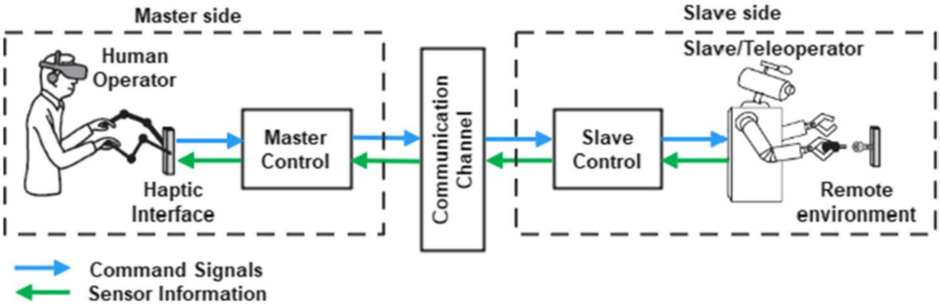


Figure 2 – A haptic bilateral teleoperation system

While having the potential to increase the operator's performance when accomplishing the task remotely, haptic bilateral teleoperation systems face two main challenges when forces are directly presented to the operator's hand/arm: the uncertainties of the human operator (dynamic, decisions) and the imperfectness of the communication channel (packet loss, delays). These problems can be overcome using force feedback surrogates instead.

[18] define sensory substitution as "the provision to the brain of information that is usually in one sensory domain (for example visual information via the eyes and visual system) by means of the receptors, pathways and brain projection, integrative and interpretative areas of another sensory system, (for example visual information through the skin and somatosensory system). Some examples include sign language for the deaf, and Braille for the blind". Ferrell [19] suggested that the advantages of force sensitivity could be maintained in the presence of a time delay if the force feedback were substituted through other modalities (auditory/tactile), and that a tactile display to the active hand might be especially compatible.

Other works explored the impact of substituting the force feedback by visual information. [20] studied the effects of visual force feedback on tying surgical knots with fine sutures like those used in cardiovascular surgery. Among surgeon subjects with robotic experience, no differences in measured performance parameters were found between robot-assisted knot ties executed with and without visual force feedback. Among surgeons without robotic experience, however, visual force feedback was associated with lower suture breakage rates, peak applied forces, and standard deviations of applied forces. Visual force feedback did not impart differences in knot completion times or loose knots for either surgeon group.

Figure 3 – Knot tying, visual Force Feedback via circle overlays. Minimal force (green circles, left), excessive applied force (red circles, right)

In the same field (surgery), [21] integrated the force information into the surgical scene by highlighting the area around the point of contact while preserving salient anatomical features. Results from the retraction and dissection studies show a clear improvement of the median forces exerted, maximum forces exerted, and time spent over the force threshold.

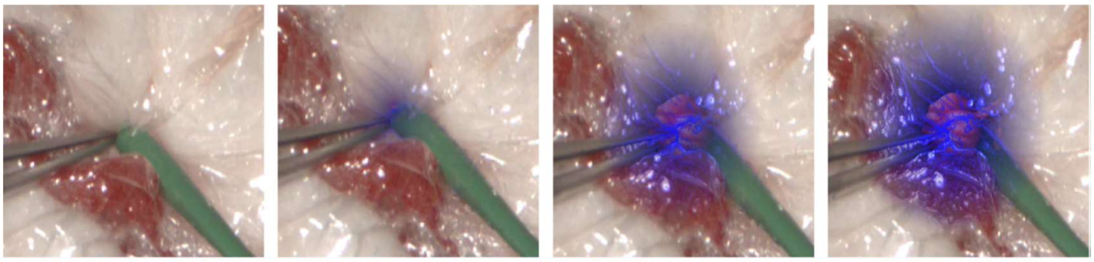


Figure 4 – Gradual increase in exerted forces

4. HAPTIC DEVICES

A haptic device is an instrument that is used to perform different kinds of manipulations with objects, providing the operator with feedback, i.e., tactile feedback, vibrational feedback, force feedback, etc.

The most widely used architectures of haptic devices are serial, parallel or hybrid. Serial haptic devices have a serial kinematic structure, which consists of a single open-loop chain of links and joints that connect the end effector to a fixed base. Parallel haptic devices have a closed-loop kinematic structure which connects the end effector on the moving platform to a fixed base with several kinematic chains, where each kinematic chain consists of passive and active parallel kinematic pairs. A hybrid kinematic device combines parallel and serial kinematic structures, where the parallel structure connects to the fixed-based, and the serial structure connects the end-effector to the parallel structure, which generates additional end-effector DOF [22]. Compared to a serial structure, a parallel mechanism has potentially larger stiffness and bandwidth as well as higher accuracy and lower inertia but provides a smaller workspace. A hybrid mechanism may combine the advantages of the serial and parallel kinematic structures.

Figure 5 – Commercial examples of haptic devices. From left to right: Sensable-Phantom Omni (serial), Quanser-HD2 (parallel), ForceDimension-Omega.7 (hybrid)

Most of the haptic interfaces available in the market are only suited for applications in which the dimensions of the task workspace do not exceed their workspace size using the typical pure one-to-one position control. Teleoperated demolition makes use of big moving slaves which range of motions clearly exceeds the haptic ones. Therefore, different control paradigms are required.

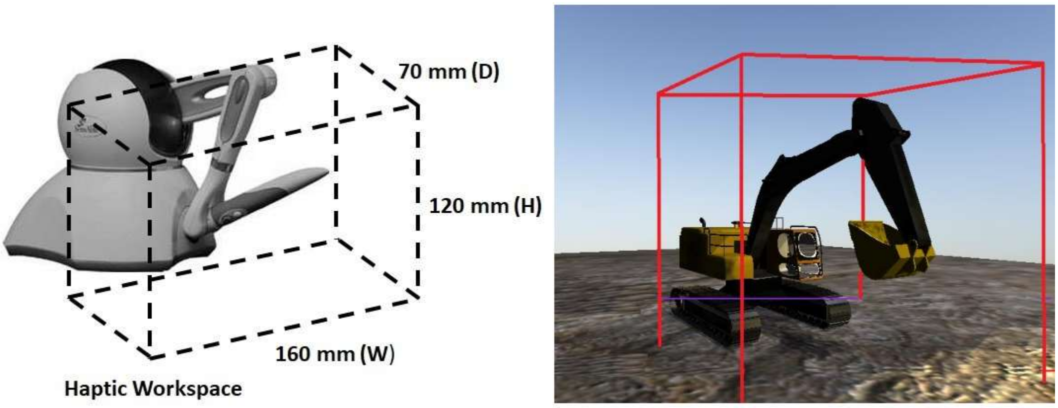


Figure 6 –Backhoe hydraulic excavator workspace is around 60 times the workspace of the Phantom Omni

The general character of commercial haptic devices makes them adaptable to almost any application. However, sometimes, none of the devices matched the requisites or was good enough for the application and efforts have been done to design new haptic interfaces.

4.1. Extending haptic workspace

There are three well-known techniques to cover large task workspaces with the limited physical workspace of a haptic device. The first one is to apply scaling in position control. It consists of amplifying the small displacements of the device to large motions of the tool to allow the operator to easily reach targets in all the task workspace. This scheme works well for coarse motion, when large distances need to be traversed to bring the tool from one area to another, but not for those in which accuracy is critical [23]. This drawback could be solved using ballistic tracking [24]. This technique consists of varying the scaling depending upon the velocity of the haptic device in its workspace. The assumption is that if the haptic is being moved very quickly, the operator is likely to be performing a coarse motion. On the contrary, if the movements are slow, then the user is likely to be performing fine positioning task. The problem arises when the offset between the frame of the haptic and the frame of the target becomes too large and the user may not able to reach some parts of the workspace. This offset can be corrected trough indexing/clutching. This technique consists of an additional switch/clutch to decouple the haptic device from the tool and allow the user to reallocate the haptic at the centre of its physical workspace [23]. It works in a similar way to a computer mouse moving on a pad. When the user reaches to the limits of the haptic (mouse reaches to the limits of the pad), he/she releases the switch (lifts the mouse), moves the haptic in the opposite direction (moves the mouse) and presses again the switch again to continue with the task (leaves again the mouse on the pad). However, indexing highly interferes with the operator because he or she needs to constantly perform it. It is also a time-consuming process which increases the total task time and which requires some training because it is not intuitive.

People do not often notice small deviations of a physical member of their body in physical space unless that small deviation has a corresponding visual deviation. [23] proposes a new control strategy (Workspace Drift Control) which make use of this observation about human perception to overcome indexing problem. Slowly shifting the workspace of the haptic device when the tool is moving instigates the user to unconsciously correct this drift while executing at the same time a task with the device. The drift is negligible when device is located near the origin, but its amplitude increases when it is on the edges of the workspace creating and important distortion between the operator's hand and the tool.

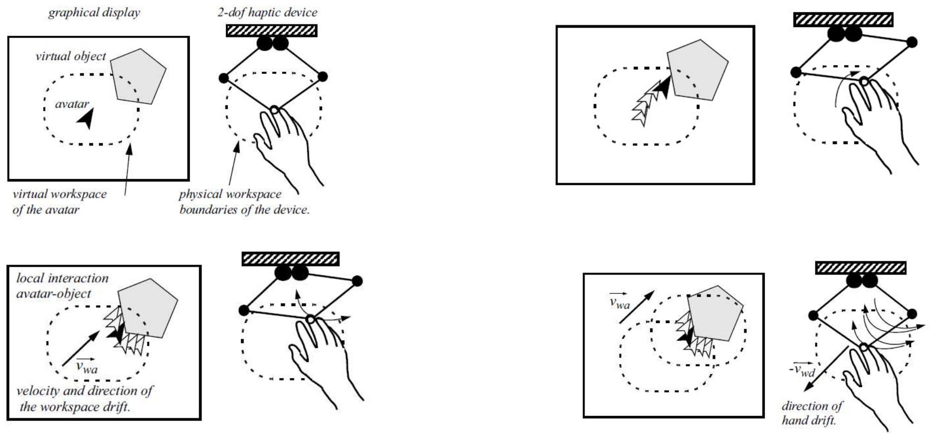


Figure 7 – Workspace drift control process.

Top-Left: the end-effector of the haptic device is in the center of its physical workspace; the physical workspace is mapped inside the task workpsace. Top-Right: The operator moves the avatar to the object reaching to the limits of the physical workspace. Bottom-Left: The operator interacts with the object. Bottom-Left: the virtual workspace drifts towards the avatar. Bottom-Right: The workspace drift instigates the operator to move his or her hand towards the center of the workspace of the device.

Other approach to overcome previous problems is the Bubble technique [25]. It uses hybrid position/rate control which enables both accurate interaction and coarse positioning in a large workspace. The haptic workspace is displayed visually using a semi-transparent sphere (looking like a bubble) that surrounds the manipulated tool. When the cursor is located inside the bubble, its motion is position controlled. When the cursor is outside, it is rate controlled. The user may also "feel" the inner surface of the bubble, since the spherical workspace is "haptically" displayed by applying an elastic force feedback when crossing the surface of the bubble.

The bubble technique presents some problems when dealing with object grasping because it focuses on one-point touching tasks. The lack of distinction between bubble and collision forces, and the fact that bubble movement may proceed even during collision, can result in an uncomfortable perceived stickiness between the colliding objects. A promising approach for addressing these issues is a non-monotonic rate control for the bubble. [26] proposes an alternative rate control function locating a peak bubble velocity a short distance outside of the position-only area of the bubble. This way the user is not encouraged to push harder/further to continually increase the velocity.

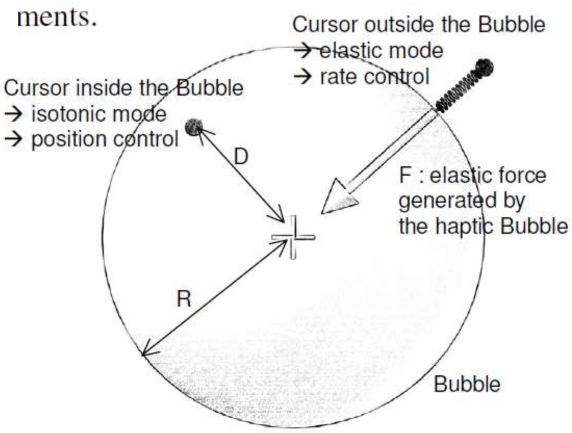


Figure 8 – Control modes in the bubble technique

4.2. Specific haptic designs for construction

The same excavators which are used to construct buildings are also utilized for dismantling processes substituting their buckets by crushers and breakers. Dismantling processes are very dangerous and, as operator must ride on the excavator, many safety-related accidents happen. Remote controlled devices are necessary to guarantee an operator's safety.

While in [27] they proposed to use a commercial haptic device (Phantom Omni) to teleoperate an excavator, in [28] they decided to develop a new one. The new haptic device was designed like the shape of excavator, so operator could understand the motion of excavator with haptic device easily: the motion of the finger controlled the bucket angle, the rotation of wrist controlled the angle of arm and arm controlled the angle of boom.

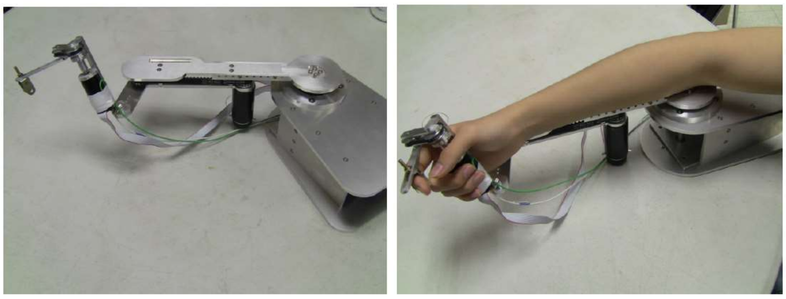


Figure 9 – New haptic interface design for the teleoperation of excavators based on their shape

5. DEMOLITION ROBOTS

In the last decade, companies such as Husqvarna, Brokk and Finmac introduced teleoperated demolition machines. These machines rely on the presence of an operator nearby. Through a wearable console, composed of a series of levers and joysticks, the operator has the possibility to manoeuvre the machine but has no feedback during the operation but visual. New interaction and control models need to be proposed to increase the safety and productivity of teleoperated demolition robots.

5.1. Localization and navigation

A previous step to the demolition task consists of placing the machine in the most suitable place for accomplishing it.

Despite its benefits (avoid risky situations for operators, reduce infrastructure execution times, improve availability of equipment, increase productivity control, improve real time-monitoring), the fully autonomous navigation of these kind of machines at a construction worksite is limited nowadays. There have been several pilots of driverless heavy construction machinery but in controlled environments. ACCIONA ([29], [30]) tried this technology on an autonomous Dumper truck at its Noblejas machinery park, where real construction site conditions (for e.g. obstacles, movement with and without load, turns that are difficult to maneuver, layout changes, route changes, etc.) were simulated.

Figure 10 – Dumper 773F. It is a common machinery in the execution of large infrastructures, as it is used for moving soil. This truck is more than 4.3 meters high and more than 10 meters long and has a load capacity of 50 tons.

However, the lack of norms/permissions to use these machines at worksite, skills availability to operate these machines and limitations on the network in the worksite (as a slight lapse could result in accident or could result in over or less excavation), limit their use in real scenarios.

An intermediate step before achieving it could be a human supervised navigation. In [31], a teleoperated navigation system based on haptic stimulation was developed to improve the awareness of the operator on the surrounding environment. The errors in positioning and navigation and the number of collisions with the environment were reduced when haptic feedback was present.

Figure 11 – Brokk demolition machine used in [31]

5.2. Cartesian control

Many companies have successfully introduced teleoperated demolition machines, however, their human interface is still primitive. Operators use levers to control each of the joints. Since these machines have multiple degrees of freedom (DoFs), operators need months to get familiar with the control system. Even with this long training, the performance of the work cannot be guaranteed because the operator does not have a preview of the result of his/her joystick manoeuvres or a situational awareness difficult to obtain from the images of a 2D camera.

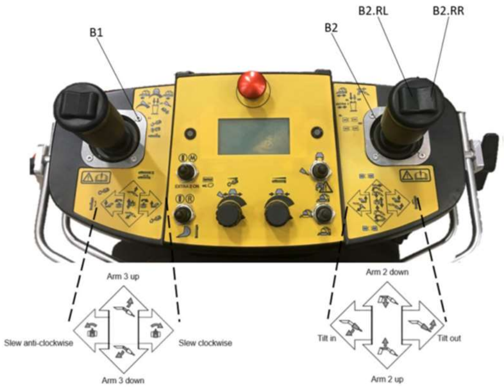


Figure 12 – Example of the interface of a teleoperated demolition machine

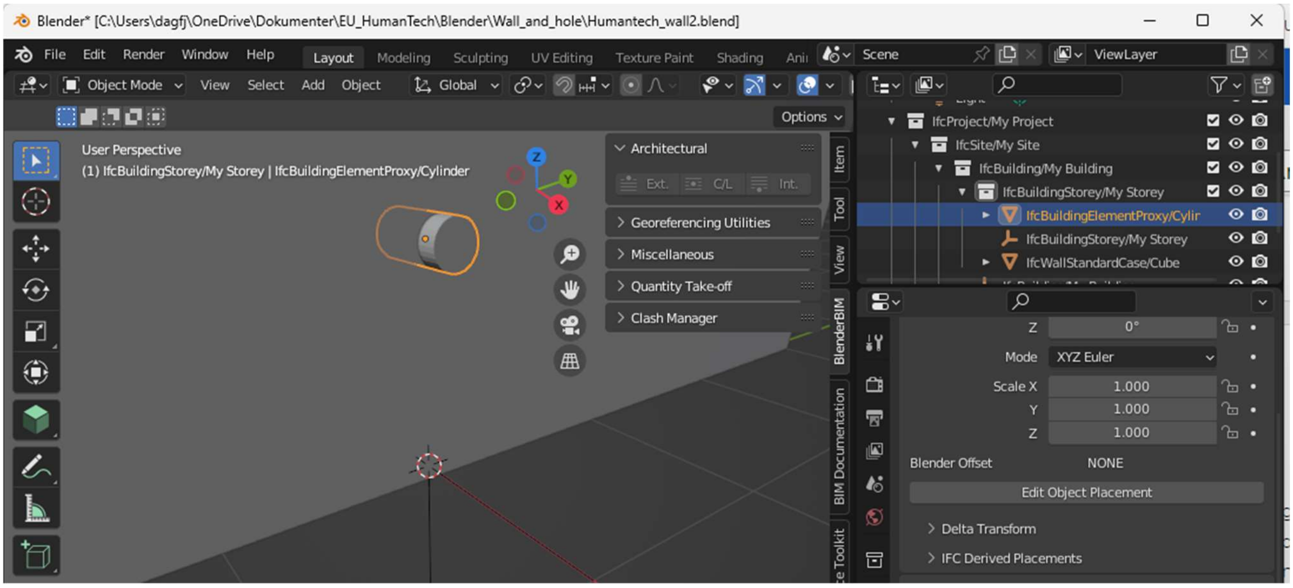
Techniques already used in traditional robotics may be of help for this kind of machines. A cartesian control of the tool could be considered a valued feature, since these machines consist of multiples DoFs and focusing only on the tool movements would reduce the task complexity. [32] presented an embedded hardware and software architecture that allow the Cartesian position control for a teleoperated construction machine. The proposed approach was implemented on an off-the-shelf demolition machine BROKK 170. The nonlinearities of the hydraulic system greatly affected the precision of the system. Although complete control of the path was not possible, reasonable results were obtained for PTP movement control.

5.3. Force/torque sensing

Scientists, engineers and psychologists have done a lot of studies related with teleoperation, and most of them agree in one point: in order to achieve good results in teleoperated systems force reflection must be used. Industrial robot arms have integrated sensors that allow to know the torque that is being applied in each of its joints or/and the cartesian wrench at their flange. This is not the case for construction machinery so to obtain the information one could add the necessary sensors to obtain the information or estimate it from other measurements. The modification of the structure could lead to a loose of strength in the structure and an increase in its price. In [33] they developed a new method to obtain the forces at the tip of the bucket of an excavator based on the incremental pressure in the chambers of the cylinders.

5.4. Connection with BIM using BCF

An innovative way to communicate for robotic demolition tasks is to create a task using BCF (BIM collaboration standard – a buildingSMART open standard for BIM oriented task collaboration). The idea is to describe both the environment and the demolition work using BIM. If the task is to create a hole in a wall, there could be a separate object, potentially in a separate model, to describe the hole that should be made geometrically and with text if needed. This “hole object” could then be the negative, meaning that the result is supposed to the situation with today’s wall with the hole subtracted.



Blender* [C:\Users\dagfj\OneDrive\Dokumenter\EU_HumanTech\Blender\Wall_and_hole\Humantech_wall2.blend]

User Perspective
(1) IfcBuildingStorey/My Storey | IfcBuildingElementProxy/Cylinder

IfcProject/My Project
IfcSite/My Site
IfcBuilding/My Building
IfcBuildingStorey/My Storey
IfcBuildingElementProxy/Cylir
IfcBuildingStorey/My Storey
IfcWallStandardCase/Cube

Z 0°
Mode XYZ Euler
Scale X 1.000
Y 1.000
Z 1.000
Blender Offset NONE

Figure 13 – Modelling the opening using BlenderBIM

The Fig.13 shows BlenderBIM used for modelling a hole that should be made in a wall. BlenderBIM is an open source BIM editing tool. The model(s) can then be exported to IFC and referred to with a BCF task. Then this task can be assigned to the robot operator. The purpose is first to communicate to the operator about the task to be done. A more advanced flow can be to use the BCF task and the related BIM models to input into the control station for the robot for technical mission planning.

6. TELEOPERATION FRAMEWORK IN HUMANTECH

The main outcome of "T5.2- Remote interfaces for demolition" consists of a teleoperation framework (operator's console, robot platform and control middleware) with advanced viewing/haptic capabilities to improve the sense of telepresence of the operator when used in hazardous tasks as demolition. This framework has been designed having in mind the applications to be developed in other tasks of the WP5 (e.g. Teaching by teleoperation in T5.5).

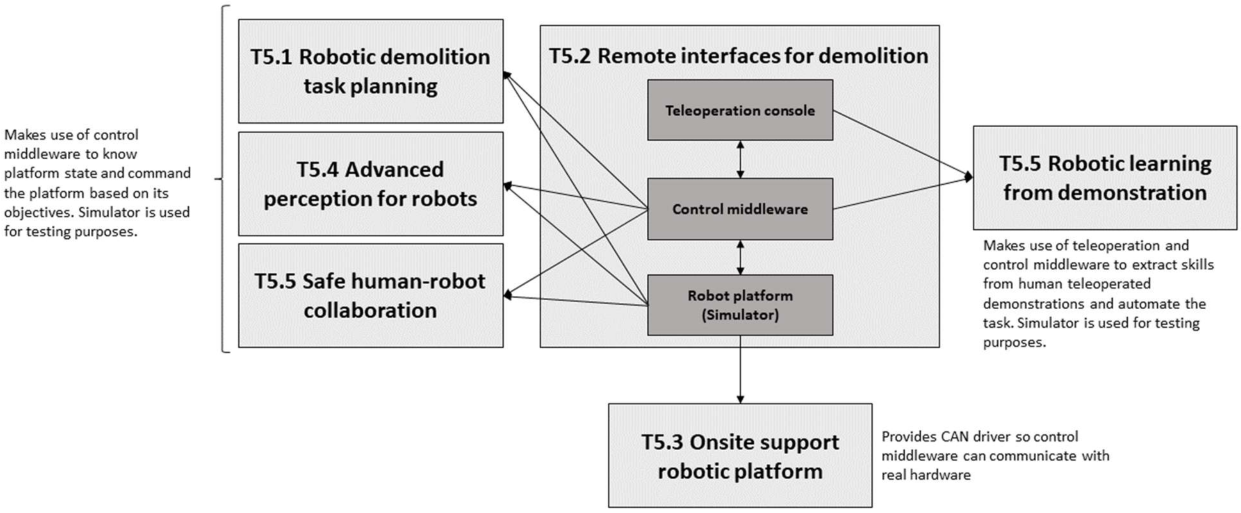


Figure 14 – WP5 tasks relationships

6.1. Operator's Console

This point describes the elements included in the operator's console that allows the operator to teleoperate the robotic arm used in HumanTech.

6.1.1. Haptic interface

The haptic interface that it is used is a Haption Virtuose 6D. This device was selected because it has a big workspace (the operator can move all his/her arm almost without touching) and the forces/torques it can transmit are big for a serial type device (peak: 35N-3.1Nm, continuous: 10N-1Nm). It has also a couple of programmable buttons which are going to be used for controlling the tools the robotic arm will hold on to its flange (i.e. mastic extruder).

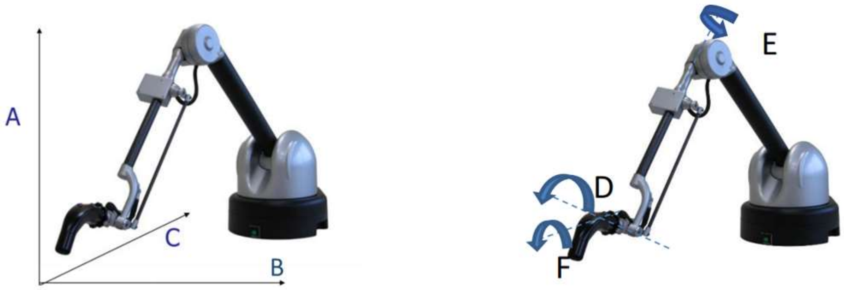


Figure 15 – Virtuose 6D haptic has 6 DoF (3 translational, 3 rotational)

Because of its intuitiveness for the operator, HumanTech has taken the approach to control the cartesian position of the tip of the tool using the haptic interface. However, the serial configuration of the Virtuose6D is similar to the one excavators have. This would allow also for a one-to-one joint control, blocking some of the rotational DoF, getting the same behaviour that it is achieved with task-specific haptic designs.

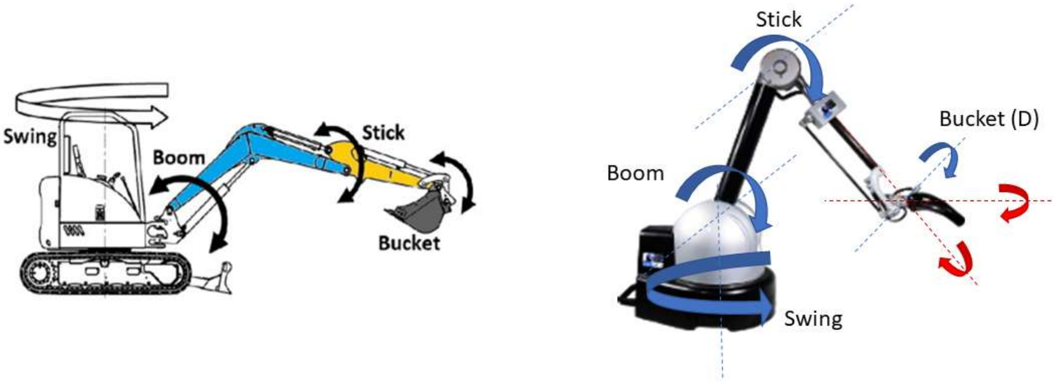


Figure 16 – Excavator DoFs and corresponding DoF in the Virtuose 6D (two of the rotational DOF, in red, should be blocked by SW or HW)

6.1.2. Footswitch

HumanTech's teleoperation console has a footswitch. It is used to engage/couple haptic movements with robot movements, so the robot arm only moves when the footswitch is pressed and the haptic moves. Although thought for security reasons, the footswitch also allows for extending the workspace of the haptic device. While the footswitch is not pressed, the operator could centre the haptic device on its workspace if he/she has reached the limits of the physical workspace. This technique is called indexing and has some problems as previously commented (interference with the task, lack of intuitiveness, etc.). This is why HumanTech has also implemented a technique like Bubble technique for extending the workspace of the haptic. The technique is called Clutchless and the main difference with the Bubble technique is a viscosity effect it has been added because instabilities were detected when changing between position/ratio modes quickly.

6.1.3. Stereo monitor

In traditional teleoperation of demolition machines, the operator was near the machine, so he/she had a direct view of the machine and the area to be dismantled. However, demolition environments are risky for the operators and teleoperation has evolved so operators can control these machines from safe “remote” places. The direct view has been removed in counterpart and substituted by an indirect view through a monitor.

Stereoscopic displays offer better depth perception and environmental awareness than standard monocular monitors [34]. Increased depth information leads to reduced collisions with the surrounding environment and better performance during highly dexterous manipulation tasks [35] (up to 40% in the case of [15]).

Because of the reasons exposed, HumanTech’s teleoperation console integrates an open passive stereoscopic display. It is necessary the use of glasses to see the 3D image, but they are lightweight and do not interfere with other equipment of the operator. It is preferable an open concept of the console because sometimes the console is not far away from the area of teleoperation and the procedure requires communication with the work team.

6.2. Robot platform

The robot platform in HumanTech consists of a commercial collaborative robot arm (Yaskawa HC20DTP) mounted on a mobile platform. The robot base is inclined with respect the mobile platform to increase the reach of the arm and try to minimize the number of configurations where the robot will be in a singularity when accomplishing manipulation tasks.

The robot platform is controlled by an on-board PLC. External computer can communicate with this PLC using CAN based communication protocol and be informed about the robot platform state and command it. The robot platform state consists of robot arm joints position, robot arm joints torque, robot arm wrench and tracks velocity. This information is taken from integrated sensors in the robot arm (encoders, torque sensors) and in the tracks (encoders). Using the same communication channel, an external PC could set the robot arm joints position and mobile platform tracks velocity.

6.2.1. Localization and navigation

One could know the position of the robot arm flange regarding its base with high accuracy thanks to the integrated sensors in the arm. However, accurate localization of the mobile platform using only the sensors integrated in the tracks is not possible.

The localization and navigation are based on the BIM model of the construction site and a 360° lidar scanner. Initially, a rough localization is done based on the geometry of the rooms similar to [36]. The walls in the current room are identified, and based on their geometry, a suitable position in the BIM model is found. Due to the uniformity of some buildings, there might be ambiguities that lead to multiple possible positions. In this case, the next visited rooms are also taken into account, and the position is estimated based on the history of visited rooms and an estimate of the movement based on the vehicle's odometry. For precise localization, the corners of the room and their distance to the vehicle are identified, see [37]. Based on those, the position and orientation are estimated using registration [38]. As a validity check, the odometry is used again to check the current position based on the previously measured one. The navigation creates target points based on the current and provides target positions for the next task. While following the path, obstacles are detected, and the planned trajectory is adjusted based on them. Since the vehicle only drives semi-autonomously, there is no dedicated safety system except the LiDAR system used for navigation. Therefore, a human has to monitor the navigation process and interfere in critical situations.

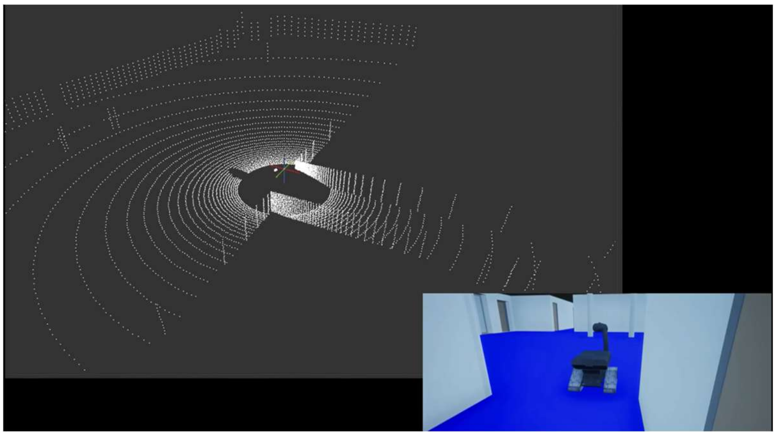


Figure 17 – Simulated laser scan in a building import from IFC file

6.2.2. On board camera

The 3D images that will be displayed in the teleoperation console will come from a stereo camera mounted on the robot platform. As the teleoperation task is more intuitive when the console (haptic) and the camera frames are aligned, the camera will be mounted on the mobile platform pointing forward instead of on the robot arm's flange (the first is static while the latest moves).

The configuration of stereo camera with wide field of view will bring enhanced awareness of the environment to manipulators. The horizontal and vertical field of view is 360 deg. and 270 deg., respectively. Stereo camera consists of two cameras mounted apart with the certain distance, which is called baseline. The cameras are time-synchronized and collect images at the same time. If an object in the environment is projected onto the images, there would be disparity between two camera images. By using the estimated disparity and known baseline, triangulation to get the distance from camera to the object is processed. With the combination between the stereo configuration and wide field of view of the camera, the manipulators can perceive the environment more accurately so that they can safely and precisely control the robot. For instance, by reflecting depth information to stereo monitor described on 6.1.3, manipulators can intuitively understand the distance between the robot and objects. The depth information with wide field of view will be utilized not only for manipulators' enhanced perception, but also for object detection by the combination with image recognition. For instance, image-based semantic segmentation detects and identify objects and creates image domains based on the instance. By associating this instance and depth information calculated by triangulation, 3D semantic environment will be reconstructed, which will lead to various kinds of application such as worker safety sensor and autonomous route guiding.

.....

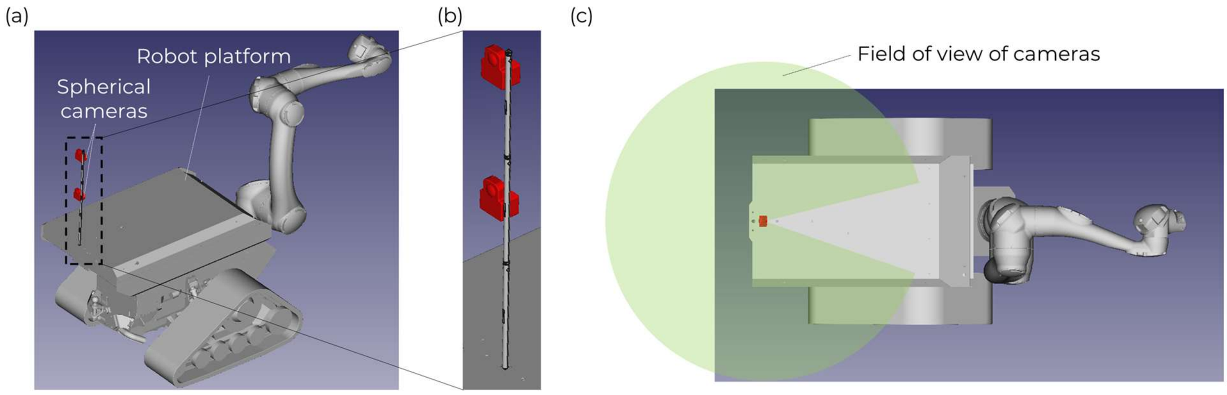


Figure 18 – (a) Spherical cameras mounted on robot platform, (b) enlarged view of the camera and (c) Top view of cameras' position

6.2.3. Simulator

A simulator of the robot platform has been developed in HumanTech. This simulator will allow partners to program and test their applications, algorithms and/or controllers without having physical access to the robot platform.

During the development of robotic systems, extensive testing is required to verify the system's functionality. In general, tests on robotic platforms are costly and time-consuming since the platform typically introduces many additional error sources, for example, due to hardware problems. Especially in the early development phase, when the availability of the hardware is not given or for experiments of high-level functionality, simulated testing is a significant advantage. The simulation of the Baubot is done using the Unreal Engine 5. The simulated robot is based on a CAD model of the Baubot and has similar functionalities. The belts and joint angles can be controlled the same way as the actual robots, and the simulation interface takes the same control commands as the Baubot. The robot's internal state is sent, and various sensor data are sent as feedback from the simulation. The communication between the engine and the control software is done via a UDP socket, which makes it independent from the used operating system.

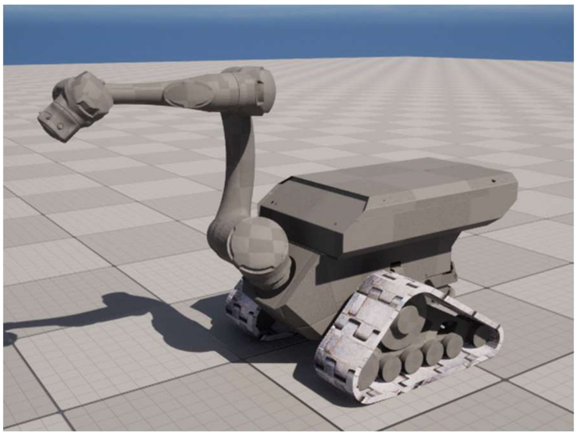


Figure 19 – Model of the robot in UnrealEngine

When accomplishing teleoperation, the operator has only the view provided by the camera mounted on-board. The simulator could be used also as part of the GUI to visualize the status of the robot platform. This will provide the operator with a wider perspective, letting him/her know how the robot is located/oriented/configured with respect the environment. It could be used also as a way to train and familiarize the operator with the teleoperation framework.

6.3. Control middleware

HumanTech's control middleware is based on the most widespread robot operating system, ROS [39], more concretely on its ros_control package [40]. This package structures the control system by means of two elements: the hardware interfaces and controllers. The hardware interfaces serve as a connection to ROS of the real elements, either sensors or actuators. They specify where the information from the sensors is positioned in memory and can be exercised on the actuators. On the other hand, controllers take on the same functions and tasks as a regulator. ros_control allows to change controller type by means of an agent called controller_manager, which allows to load, delete, or change controllers in real time.

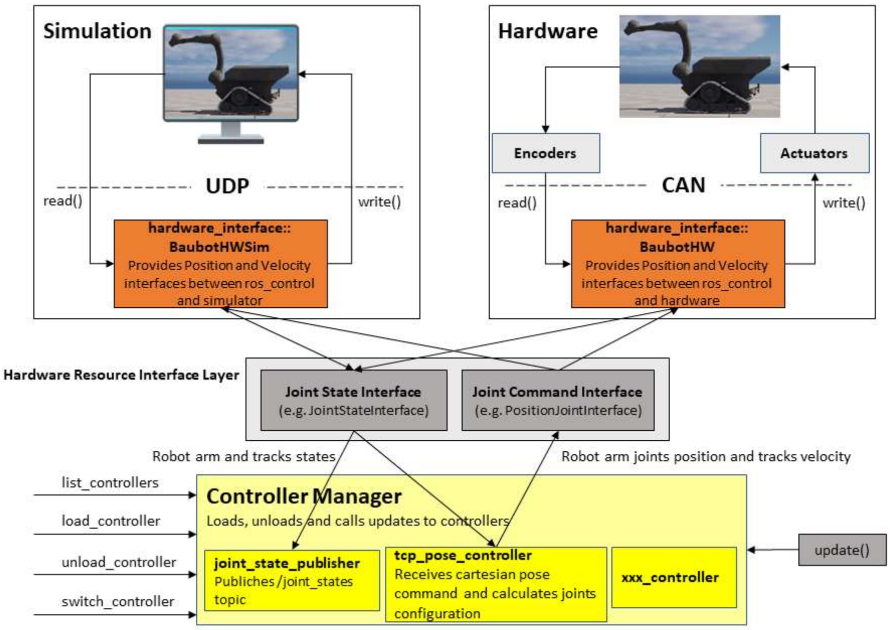


Figure 20 – HumanTech's Control Middleware

The control middleware in HumanTech has been designed in such a way the environment with which the algorithms and controllers interact (real hardware or simulation) can be changed easily by editing text files (without having to build them). This allows for rapidly moving from simulation to real hardware and vice versa, reducing the development times and minimizing the risks when testing these developments on the real platform.

Some controllers have been integrated in the control middleware. These controllers expose ROS topics so other "3rd party" applications can subscribe to them to listen to platform state and command the position of the robot arm joints and/or the velocity of platform tracks:

- /joint_state_publisher: it reads the state of the platform and publishes the information in a topic called /joint_state.
- /arm/position_controller: it publishes a topic to which other applications can subscribe and send the desired robot arm joints position.
- /caterpillar/velocity_controller: it publishes a topic to which other applications can subscribe to and send the desired tracks velocity.
- /arm/tcp_controller: it publishes a topic to which other applications can subscribe to and send the desired cartesian poses of the TCP of the robot arm. The controller is on charge of transforming the cartesian poses to robot joints configurations which will be written to the hardware interface.
- /extended_haptic_controller: it allows to control the haptic interface extending the physical workspace of the device. It also publishes the haptic movements which serves as input to other applications and/or controllers (e.g. admittance_teleop_controller).
- /admittance_teleop_controller: it controls the robot arm configuration based on the haptic movements received from the topic /extended_haptic_controller publishes and the forces read from the TCP of the robot arm.

The control middleware runs on Ubuntu 20.04 with ROS noetic version installed.

6.3.1. Teleoperation controller

The teleoperation controller in HumanTech is an admittance controller which has been developed based on an existing configurable constraint controller. The controller listens to haptic movements and commands the robot accordingly taking into account the force that is being applied against the environment. An approach of the applied force is also transmitted to the haptic interface, so the operator has an idea of the exerted force. This teleoperation controller has been integrated in the control middleware and it is used by T.5.5 applications to automate processes based on demonstrations done by human operators using teleoperation.

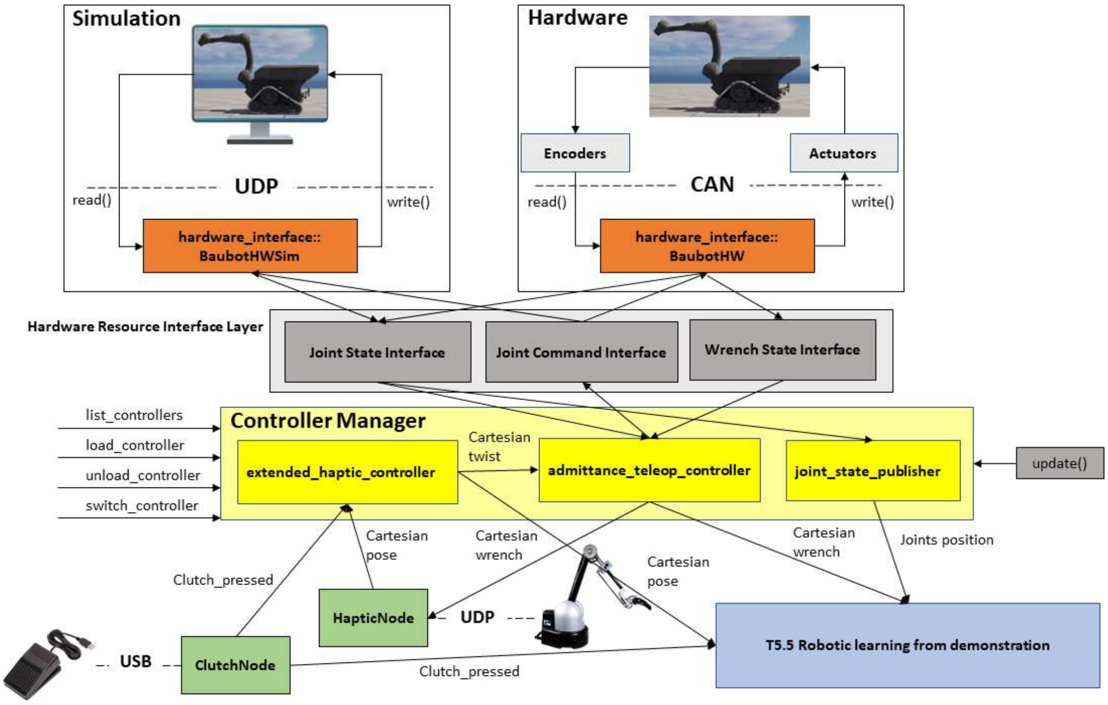


Figure 21 – Teleoperation controller in HumanTech's Control Middleware

Firstly, a position-based teleoperation controller was implemented and tested in both simulation and UR10e. While the robot arm followed the haptic movements precisely, security stops arise due when excessive forces were applied against the environment. This issue was corrected introducing the external forces perceived by the TCP of the robot arm in the loop together with a proportional gain, to tell how much the robot arm can be separated from the desired pose, and a differential gain, to smoothen the situations when big differences in force happen (first contact, collisions, etc.).

To provide the operator with an idea of the exerted force against the environment, the measurements of the force/torque sensor of the UR10e were transmitted to the haptic interface (with a proportional gain to adapt force sensing/displaying ranges). The provided forces were not stable/precise which led to an unstable control loop. So, it was decided to provide the sensation using the difference between the desired and current robot poses.

As previously commented, the admittance teleoperation controller developed in HumanTech is based on an existing configurable constraint controller. This is a ros_control controller which can solve certain types of tasks specified using configuration only. A task, in this context, is specified by defining a set of constraints and a solver. Constraints act in constraint velocity space by defining both: a Jacobian matrix, mapping joint velocity space to it’s constraint velocity space; and a desired constraint velocity. Likewise, solvers aggregate all constraint Jacobians and desired velocities and compute the joint velocities that satisfy the constraints.

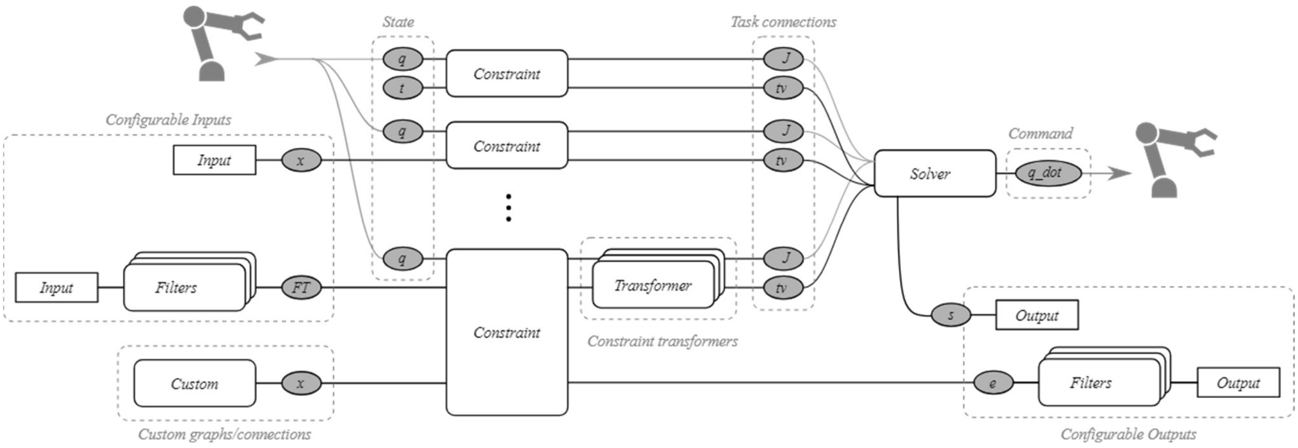


Figure 22 – Configurable constraint controller framework

Both constraints and solvers (as well as other types of graphs discussed below) are implemented as dataflow graphs and are loaded at the controller initialization time. The image below illustrates the generated graph for the constraint used for the admittance teleoperation controller.

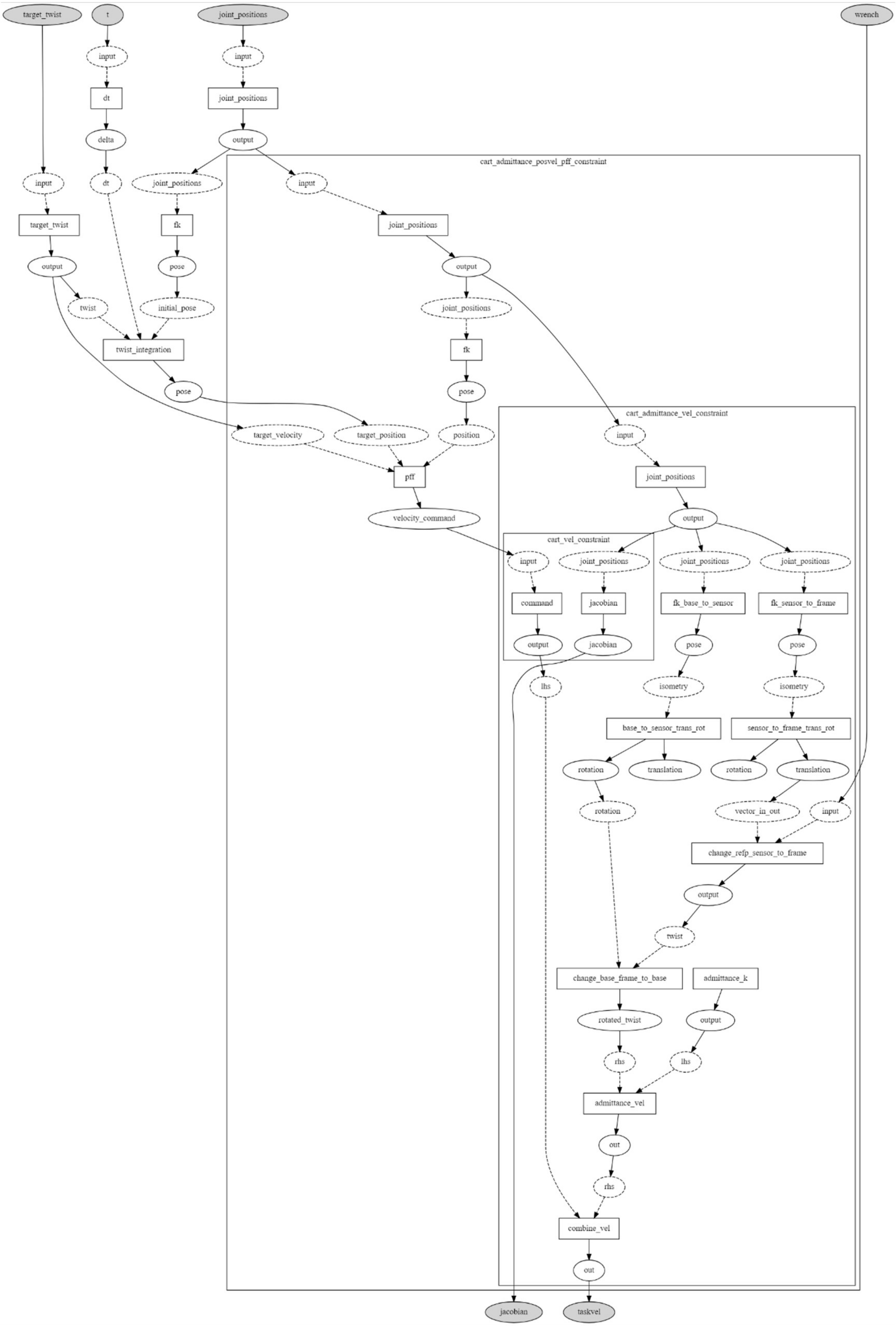


Figure 23 – Admittance teleoperation controller constraint dataflow

6.3.2. Task planning for demolition

In HumanTech a task planner for the automatic execution of demolition activities is being developed. Specifically, the demolition activity taken into consideration is the cutting of openings into existing walls, a common task during renovation projects. These openings may serve for example as conduits for essential elements such as MEP (mechanical, electrical, and plumbing) ducts and pipes, as well as additional doors or windows. Traditionally, these tasks are labour-intensive and potentially hazardous, involving human-operated tools like grinders and rudimentary guidance systems, leaving the human workers exposed to harmful environmental factors like dust and falling debris.

The concept behind the task planner is to empower HumanTech demolition robots to autonomously perform some of those critical operations, thereby mitigating the risks associated with manual labour. In this innovative approach, human operators assume a supervisory role, remotely monitoring and ensuring the task's successful execution. For the tasks for which automation is not possible, because for example real-time knowledge of a skilled operator is needed, teleoperation will be the alternative to keep him/her away from the risky environment.

The task planner outlines a series of high-level robotic activities that must be executed to accomplish the precise cutting of an opening into a wall:

- **Navigation:** the mobile platform is tasked with navigating to a pre-defined position in front of the wall. This navigation process involves several sub-tasks, including localization, path-planning, and motion control. Precise coordination is crucial to ensure the robot's arrival at the correct location.
- **Marking:** once in position, the manipulator, mounted on the mobile platform, employs a marking tool to delineate the precise position on the wall where the opening is to be made. Accurate localization is vital to ensure the marking aligns perfectly with the desired location.
- **Drilling:** following successful marking, the robot utilizes a specialized drill end-effector to perforate the wall, creating the initial opening. The drilling process demands tight control over both the tool and the robot to guarantee precision and safety.

- **Cutting:** with the preliminary hole in place, a circular sawblade is employed to execute the final cut through the wall, creating the designated opening. Again, the control of both the tool and the robot is paramount during this critical phase.

All three primary actions – marking, drilling, and cutting – necessitate the synchronized operation of the specific tool and the robot. The sequential order of these activities is a crucial point, as they must follow the precise sequence just described to ensure the successful creation of the wall opening.

Automatic task planning in the HumanTech system relies on the HumanTech ontology [41], a structured knowledge framework containing information about building demolition and the demolition environment. This ontology equips the robot with an understanding of its surroundings and the potential consequences of its actions.

This ontology is an extension of ifcOWL [42], a Web Ontology Language (OWL) representation of the Industry Foundation Classes (IFC) schema, made available by Building Smart [43]. To enhance its applicability to demolition activities, ifcOWL has been expanded with additional concepts, properties, and relationships. Specifically, new classes have been defined, like "Robot", "Tool", and "Opening", while the already existing “IfcWall” class has been extended with supplementary properties.

For instance, within the ontology:

- The "Robot" class has sub-classes "MobileRobot" and "StationaryRobot" and includes properties such as "hasRobotName", "hasTool", and "isAvailable".
- The "hasTool" property establishes a connection between a robot and the specific “Tool” it employs, which could be a "Marker", a "Driller", or a "Sawblade".
- The "isAvailable" property is represented by a Boolean value indicating whether a particular robot is ready to perform a task.

Reasoning on all this information, it is possible to find out if a robot equipped with the correct tool is at the current moment available to perform one of the demolition tasks. If it is available, the task planner can then assign the robot the task and trigger all the requisite sub-tasks seamlessly.

7. CONCLUSIONS AND FUTURE STEPS

The working environment of demolition is known as one of the most hazardous workplaces which has a high probability of fatal injuries of workers. Its full automation would be the solution for improving the security and productivity of the tasks involved in it. However, demolition robots are not fully autonomous primarily due to the unique challenges and complexities associated with the demolition environment (unpredictable and dynamic environment, complexity of the tasks, safety concern, sensory limitations, etc.). HumanTech proposes converting demolition tasks in “remotely supervised, semi-automated” processes where the robot will operate autonomously but supervised by the human operator which would intervene through teleoperation whenever it is needed.

The Human-Machine Interface (HMI) is a critical component of teleoperation systems in demolition tasks, and its design can significantly impact the effectiveness and safety of remote operations. Providing operators with a comprehensive understanding of the demolition environment, including depth perception and spatial awareness, is crucial for making informed decisions. Haptic feedback, in addition to visual feedback, plays a crucial role in teleoperation by providing a tactile sense of the remote environment and enhancing the operator's situational awareness. The importance of haptic feedback in teleoperation lies in its ability to enhance spatial awareness, improve object manipulation, increase safety, and provide operators with a more immersive and intuitive experience.

This document has provided an overview of the current state-of-the-art on haptic-visual feedback mechanisms for efficient teleoperation of demolition robots with the idea of integrating some of them in the console implemented in task “T5.2 - Remote interfaces for demolition”. This console is a part of the teleoperation framework developed in HumanTech which allows an operator to remotely operate a robot arm and interact with the distant environment in a seamless and immersive way:

- The stereo configuration of the wide field of view cameras together with the 3D display enable the operator to perceive depth accurately.
- The haptic device, used for commanding the robot, also provides the haptic feedback which allows the operator to feel the force exerted to the environment. Proper mechanisms for the extension of the haptic interface workspace have

been implemented so its use is simpler and more intuitive while keeping accuracy.

- The admittance teleoperation controller enables the remote control of the robot arm while allowing a compliant interaction of the robot arm with the environment.
- The simulator together with the control middleware allows for testing any algorithm/application before setting it up in the real robotic platform. They can be used also for user familiarization and training on teleoperation.

HumanTech’s teleoperation framework will be used among others in task “T5.5 - Robotic learning from demonstration”. In this task, methods of "teaching by demonstration" are being developed to facilitate the implementation of automated activities without explicit coding of robot programs. In this case, demonstrations are implemented through teleoperation to allow programming of heavy-duty construction robots from a safe distance. Related to this task, there will be a pilot (Pilot V [44]) consisting of automating the application of mastic material to expansion joints by learning the procedure from demonstrations done through the developed teleoperation framework.

8. REFERENCES

[1] M. Shahbazi, S. F. Atashzar and R. V. Patel, “A Systematic Review of Multilateral Teleoperation Systems,” *IEEE Transactions on Haptics,* vol. 11, no. 3, pp. 338-356, 2018.

[2] P. F. Hokayem and M. W. Spong, “Bilateral teleoperation: An historical survey,” *Automatica,* vol. 42, no. 12, pp. 2035-2057, 2006.

[3] J. Chen, E. Haas and M. Barnes, “Human performance issues and user interface design for teleoperated robots,” *IEEE Transactions on Systems Man and Cybernetics Part C (Applications and Reviews),* 2007.

[4] G. Burdea, “Haptic Feedback for Virtual Reality,” in *Proceedings of International Workshop on Virtual prototyping, La*, Laval (France), 1999.

[5] C. E. Lathan and M. Tracey, “The Effects of Operator Spatial Perception and Sensory Feedback on Human-Robot Teleoperation Performance,” *Presence: Teleoperators and Virtual Environments,* vol. 11, no. 4, p. 368–377, 2002.

[6] G. V. Popescu, G. C. Burdea and H. Trefftz, “Multimodal Interaction Modeling,” in *Handbook of virtual environments: Design, implementation, and applications,* 2002, p. 435–454.

[7] M. Mori, F. MacDorman and N. Kageki, “The Uncanny Valley [From the Field],” *IEEE Robotics & Automation Magazine,* vol. 19, no. 2, pp. 98-100, 2012.

[8] C. C. Berger, M. Gonzalez-Franco, E. Ofek and K. Hinckley, “The uncanny valley of haptics,” *Science Robotics,* vol. 3, no. 17, 2018.

[9] J. Berg and S. Lu, “Review of Interfaces for Industrial Human-Robot Interaction,” *Robotics in Manufacturing,* pp. 27-42, 2020.

[10] H. Liu and L. Wang, “Gesture recognition for human-robot collaboration: A review,” *International Journal of Industrian Ergonomics,* pp. 355-367, 2018.

[11] K. Qian, Y. H. and J. Niu, “Developing a gesture based remote human-robot interaction system using kinect,” *International Journal of Smart Home,* pp. 203-208, 2013.

[12] H. Wu, H. Li, H.-L. Chi, Z. Peng, S. Chang and Y. Wu, “Thermal image-based hand gesture recognition for worker-robot collaboration in the construction industry: A feasible study,” *Advanced Engineering Informatics,* 2023.

[13] M. L. Heilig, “El Cine del Futuro: The Cinema of the Future,” *Presence: Teleoperators and Virtual Environments,* vol. 1, no. 3, p. 279–294, 1992.

[14] R. J. Moreno, F. Espinosa and D. A. Hurtado, “Teleoperated systems: A perspective on telesurgery applications,” *Revista Ingeniería Biomédica,* 2013.

[15] E. Triantafyllidis, C. Mcgreavy, J. Gu and Z. Li, “Study of multimodal interfaces and the improvements on teleoperation,” *IEEE Access,* vol. 8, pp. 78213-78227, 2020.

[16] R. L. Thompson, “Integration of Visual and Haptic for Teleoperation,” 2001.

[17] S. Nahri, S. Du and B. Van Wyk, “Review on Haptic Bilateral Teleoperation Systems,” *Intelligent & Robotic Systems,* vol. 104, no. 13, 2022.

[18] P. Bach-Y-rita, J. G. Webster, W. J. Tompkins and T. Crabb, “Sensory Substitution for Space Gloves and for Space Robots,” in *Proceedings of the Workshop on Space Telerobotics*, 1987.

[19] W. R. Ferrell, “Delayed Force Feedback,” *Human Factors,* vol. 8, no. 5, pp. 449-455, 1966.

[20] C. E. Reiley, A. T., D. Burschka, D. C. Chang, A. M. Okamura and D. D. Yuh, “Effects of visual force feedback on robot-assisted surgical task performance,” *Thorac Cardiovasc Surg,* vol. 135, no. 1, pp. 196-202, 2008.

[21] G. Gras, H. J. Marcus, C. J. Payne, P. Pratt and G. Z. Yang, “Visual Force Feedback for Hand-Held Microsurgical Instruments,” in *Medical Image Computing and Computer-Assisted Intervention*, Munich (Germany), 2015.

[22] X. Sun, K. Andersson and U. Sellgren, “Design optimization of haptic device - A systematic literature review,” 2017.

[23] F. Conti and O. Khatib, “Spanning large workspaces using small haptic devices,” in *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics Conference*, Pisa (Italy), 2005.

[24] k. Salisbury, F. Conti and F. Barbagli, “Haptic rendering: introductory concepts,” *IEEE Computer Graphics and Applications,* vol. 24, no. 2, pp. 24-32, 2004.

[25] L. Dominjon, A. Lecuyer, J. M. Burkhardt, G. Andrade-Barroso and S. Richir, “The "Bubble" technique: interacting with large virtual environments using haptic devices with limited workspace,” in *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics Conference*, Pisa (Italy), 2005.

[26] R. A. Pavlik and J. M. Vance, “Expanding haptic workspace for coupled object manipulation,” in *World Conference on Innovative Virtual Reality* , Milan (Italy), 2011.

[27] J. Frankel, “Development of a Haptic Backhoe Testbed.,” 2004.

[28] D. Kim, K. W. Oh, N. Kim, B. Chu, D. Hong, Y. K. Kim and S. Hong, “Design of haptic device for excavator with pressure transmission,” in *International Symposium on Automation and Robotics in Construction*, 2010.

[29] “Driverless construction machinery,” [Online]. Available: https://www.youtube.com/watch?v=855J7eOU69s. [Accessed 16 10 2023].

[30] “ACCIONA news website,” [Online]. Available: https://www.acciona.com/updates/news/acciona-conducts-first-trials-driverless-public-works-machinery-toledo-depot/?_adin=02021864894. [Accessed 15 10 2023].

[31] S. Patrinostro, M. Tanzini, M. Satler, E. Ruffaldi, A. Filippeschi and C. A. Avizzano, “A Haptic-Assisted Guidance System for working machines based on virtual force fields,” in *2015 XXV International Conference on Information, Communication and Automation Technologies (ICAT)*, Sarajevo (Bosnia and Herzegovina), 2015.

[32] H. Lee and S. Brell-Cokcan, “Cartesian coordinate control for teleoperated construction machines,” in *Construction Robotics*, 2021.

[33] A. Barrientos, R. Aracil, O. Luengo and E. Gambao, “The Teleoperation of a Backhoe Excavator through Force Reflection and Pressure Measurements.,” 1997.

[34] L. B. Rosenberg, “The effect of interocular distance upon operator performance using stereoscopic displays to perform virtual depth tasks,” in *Virtual Reality Annual International Symposium*, Seattle (USA), 1993.

[35] D. R. Scribner and J. W. Gombash, “The Effect of Stereoscopic and Wide Field of View Conditions on Teleoperator Performance,” 1998.

[36] H. Blum, J. Stiefel, C. Cadena, R. Siegwart and A. Gawel, “Precise Robot Localization in Architectural 3D Plans.,” 2021.

[37] M. Ahmed, Y. Tan, C. Chew, A. Al-Mamun and F. Wong, “Edge and corner detection for unorganized 3d point clouds with application to robotic welding,” *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),* 2018.

[38] X. Huang, G. Mei, J. Zhang and R. Abbas, “A comprehensive survey on point cloud registration,” 2021.

[39] “ROS - Robot Operating System,” [Online]. Available: https://www.ros.org/. [Accessed 20 09 2023].

[40] “ros_control package,” [Online]. Available: https://wiki.ros.org/ros_control. [Accessed 20 09 2023].

[41] “D5.1 - Development of the ontology for demolition task planning (T5.1 - Robotic demolition task planning)”.

[42] “ifcOWL,” [Online]. Available: https://technical.buildingsmart.org/standards/ifc/ifc-formats/ifcowl/. [Accessed 5 10 2023].

[43] “BuildingSMART,” [Online]. Available: https://technical.buildingsmart.org/. [Accessed 05 10 2023].

[44] “D1.2 - HumanTech architecture definitions (initial) (T1.3 - Architecture definitions)”.