Grant Agreement N°101092861

Making the Invisible Visible for Off-Highway Machinery by Conveying Extended Reality Technologies

DELIVERABLE 4.3 – INTERACTION SEQUENCE MODELS (FIRST VERSION)

This work is part of the THEIA^{XR} project that received funding from the European Union's Horizon Europe research and innovation program under grant agreement no. 101092861

Funded by the European Union

Document Identification

Project	THEIA^{XR}
Project Full Title	Making the Invisible Visible for Off-Highway Machinery by Conveying Extended Reality Technologies
Project Number	101092861
Starting Date	January 1st, 2023
Duration	3 years
Horizon Europe Programme	Horizon Europe Framework Program (HORIZON)
Topic	HORIZON-CL4-2022-HUMAN-01-14: “eXtended Reality Technologies”
Call for proposal	HORIZON-CL4-2022-HUMAN-01
Type of Action	RIA-Research & Innovation Action
Website	www.theia-xr.eu
Work Package	WP4
WP Leader	TUD
Responsible Partner(s)	HdM
File Name	D 4.3 – Interaction sequence models (first version)
Contractual delivery date	30.11.2023
Actual delivery date	22.12.2023
Version	1
Status	Final
Dissemination level	Public
Author(s)	Manuel Kulzer (HdM), Sebastian Lorenz (TUD), Kaj Helin (VTT), Thomas Howard (HAP), Martijn Rooker (TTC)
Contact details of the coordinator	Martijn Rooker, martijn.rooker@tttech.com
Internal Reviewer	Daniel Roeck, PRIN Martijn Rooker, TTC

Disclaimer

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.

Executive summary

In Task T4.3, the findings of user research and concepts developed within the transdisciplinary co-design approach (WP3) are used to create technical models for a cyber-physical system (CPS) and its human-machine interface (HMI). The process involves translating visions and planned interaction flows into CPS specifications. The presented interaction sequence models, based on Norman's Human Action Cycle, focus on information retrieval and the integration of XR technology to enhance vehicle operators' awareness and performance. The models emphasize improving information design and retrieval processes, moving away from traditional small screens. The conceptual models represent early solutions, pending validation in co-design workshops.

For UC1, snow grooming, three key scenarios are discussed, addressing challenges such as diverting the operator's attention during adjustments, constant checks for obstacles affecting attention, and the safety risks associated with reduced visibility in adverse weather conditions. The proposed visionary solutions involve real-time projection of information onto the snow surface, visualizing obstacles with lights, and projecting map data onto the windshield to enhance navigation in challenging conditions.

For UC2, logistics, three key scenarios were modelled: Picking up containers, placing containers, and collision avoidance during transport. Challenges include visually estimating precise spreader positioning and frequent manual checks for obstacles. Proposed solutions involve using XR projections to aid spreader positioning and employing thermal cameras to detect and highlight objects in the operator's field of view.

For UC3, construction, three key scenarios are discussed. These scenarios include achieving precise depth control during grading, obstacle and person detection, and streamlining the infield design process for excavation pits. Challenges included constant depth checking during grading, limited visibility behind the boom, and a time-consuming calibration process in infield design. Proposed solutions involve visualizing bucket depth with LED or laser projection, using ambient lights for better detection, and implementing LiDAR in combination with laser projections for accurate excavation pit outlines.

In the final chapter, two XR interface design approaches currently in development for controlling off-highway machinery are presented: One utilizing Lidar sensors and the Varjo XR-3, and another employing 360° cameras with augmented content. Furthermore, two roles of haptic feedback — informative and assistive — are discussed in optimizing operator tasks and possible applications in all use-cases are outlined.

The interaction sequence models provide a formalized framework for describing XR interaction within machine operation tasks, considering human perception and processing capabilities. These models serve as a tool for multi-disciplinary discussions. This foundational work lays the groundwork for subsequent technical development and integration within the project.

Table of Contents

Table of Figures

1 Introduction

In Task T4.3, the broad visions, information components and interaction flows described and currently in development in the transdisciplinary scenario-based co-design approach in WP3 will be translated into more technical models to derive specifications for the cyber-physical system (CPS) in general and the human-machine interface (HMI). This will also be an iterative process, as the first version presented in this deliverable will be outlining the status of the visions (activity scenarios) from task T3.2 and their implications for the main use-case scenarios, while future versions will be more detailed, following the co-design and validation steps of information scenarios and interaction scenarios.

The progress of the transdisciplinary scenario-based co-design approach (T3.2) and its translation into interaction sequence models in this task, just like further HMI design steps, is dependent on the completion of user research, specifically the contextual inquiries described in Deliverable D3.1. This user research allows us to understand the workflows and thought processes of vehicle operators in detail and enables us to co-create visions, information and interaction concepts and their respective interaction sequence models based on reality, not assumptions. As we first mentioned in Deliverable D3.1, there have been organizational problems and several delays in the user research activities for UC2 (logistics), which has resulted in a final user research study taking place at KALMAR facilities in Sweden in December 2023 instead of November. This meant that the first version of this deliverable, which was completed and reviewed by the end of November, had only included the interaction sequence models for UC1 and UC3. By delaying the submission of this deliverable by one month, we were able to integrate the results of our user research visit to KALMAR in Sweden, where the status quo of UC2 reachstacker operation was studied and first ideas for solutions were gathered.

In the following chapters, the theoretical background and our adaptation of the interaction sequence models will be described. We will then present interaction sequence models for all three use-cases, outlining the initial (problematic) situation and possible (conceptual) solutions for three use-case scenarios each. Finally, the further use of these models in the HMI design and development process will be discussed.

2 Interaction sequence models

2.1 Model background and structure

The development and integration of XR technology into operation environments for machine operators aims at improving information transfer, perception and processing on the operators' side. XR technology overlays information, available in the physical world, with information of digital origin. This advanced information coverage aims at improving the operators' situation awareness, process and machine understanding and work performance. To describe our XR technology approaches in this context in a formalized way we use interaction sequence models.

The interaction sequence models are based on Norman's "Human Action Cycle" (aka "seven stages of action", see Figure 1) [1][2], a commonly used model in the field of HMI design. His original cycle of interaction stages consists of forming an intention, selecting an action, executing the action and evaluating the outcome [1]. Execution and evaluation can in turn be divided into multiple phases [3]. We adapt and reinterpret this model to focus on the use of onboard information systems and visualisations where the action relates to the information retrieval and the environment/world (see Figure 1) relates to the information systems, e.g., screens and projections. We specifically include the perception and interpretation of outcomes (information) as we consider these to be relevant stages in the context of making previously non-perceivable information perceivable and making information easier to access and process.

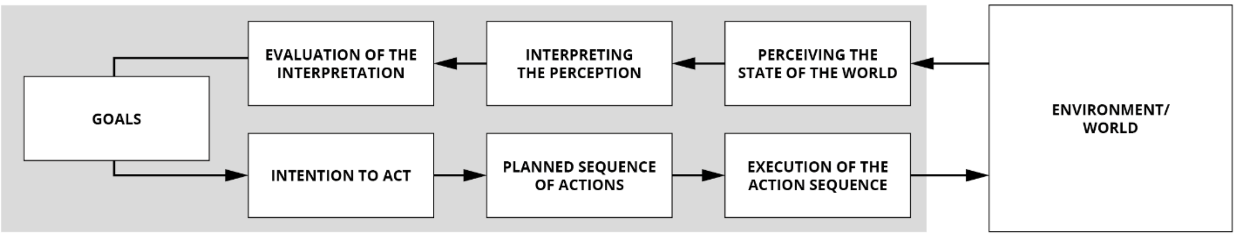

Figure 1: Seven stages of action

Our models' starting point is the general intent (stage 1), which is the respective task in each use case scenario. Based on the intent, the operator selects an action. While there are usually many actions involved in the process of completing their tasks, we focus our models on the action of information retrieval, not on the specific actions required with the machine controls to attain the goal of each use case scenario. We learned about those further actions during user research studies, but the vehicle controls themselves are not the main subject of research and development. Today, information is usually displayed on (and retrieved from) small system screens on either side of the cabin (see Figure 2). Since looking at these screens takes away attention from the main operating area in virtually all off-highway vehicles (apart from remote desk operated machines, e.g., straddle carriers), this is one of the main aspects we want to improve with XR technologies in this project.

The action of information retrieval (stage 2) in our model is followed by perception (stage 3), where the modalities of information conveyance come into play. Perception is followed by the interpretation of information (stage 4), which may today be made difficult by abstract visualizations that require conceptual translation into concrete implications for the real-world scenario. This is another main aspect we plan to improve in this project with XR technologies and more tangible information design.

The final step of our models is the action sequence and evaluation (stage 5), which comprises all following actions using the vehicle controls, e.g., throttle, steering and tool controls, which expanding on in these models would go too far into detail and is not required to understand the main improvements that can be made to information retrieval, conveyance and design. When a certain sequence of actions is completed and evaluated (e.g., adjusting the snow-groomer blade and tiller to adapt to changing snow conditions), there may be a loop back to the action stage, as continuous checks of information may be required to adapt to the new situation of the vehicle. We do not loop back to the intent as in the original human action cycle since we consider the intent of task completion to be ongoing. Beyond these five stages of interaction, we include implications resulting from the action (information retrieval) and interpretation stages.

Figure 2: Examples for off-highway vehicle cabin screens (left: excavator; right: snow groomer)

We created models in this structure for both the existing (initial) and the future (conceptual) operating situations to show the (partially problematic) status quo and planned XR and information design improvements. These initial models are based on user research. The conceptual models are based on early versions of activity scenarios created in Task T3.2, which are yet to be validated and expanded in co-design workshops with operators. Thus, the conceptual models represent ideas of possible solutions which are being explored or are in early stages of development, but their final implementation is not yet definite.

2.2 Use case 1 interaction sequence models

In use case 1, snow grooming, we modelled the initial situations and visionary interaction flows of the following three scenarios:

1. Build slope matching target surface
2. Collision avoidance
3. Zero-sight navigation

These scenarios were subject to user research in Sterzing, Italy, in March 2023 within Task T3.1.

Scenario 1 presents the main job of snow groomer operators on the slope: Building the slope to match the required target surface, which involves preparing a smooth surface, maintaining slope borders, adding and removing snow, moving excess snow to other parts of the slope and avoiding damages to the surface (e.g., accidently ripping the ground open with the blade, causing snow and soil to mix, or using a wrong configuration of the blade wings, causing chunks of snow to escape and possibly damage the already finished parts of the slope). In the model below (see Figure 4), we focus on the main task of matching the target surface, which requires knowledge of current and target snow depth. Today, this is usually handled with an onboard system that shows the delta between current and target snow depth, visualized with a heatmap where specific colours indicate a lack or excess of snow (see Figure 3). Checking the delta and adjusting the blade accordingly requires repeated checking of the map data on a screen, which may be on the side or near the bottom of the cabin, thus, taking away the operator’s attention from the operating area in front. Adjustment of the tiller is depending on snow conditions and needs to be done based on the operator’s knowledge and experience.

Figure 3: Snow-groomer cabin view. Snow depth measurement system screen in the bottom middle.

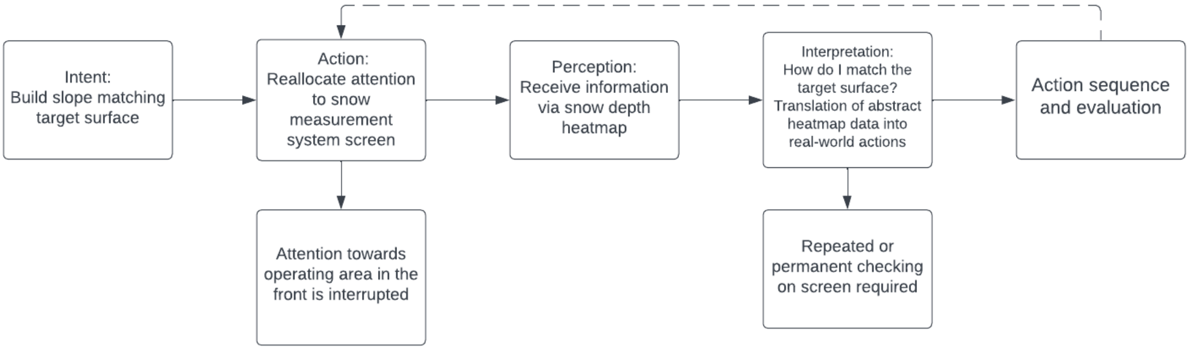

Figure 4: Initial interaction sequence model for UC1 scenario 1.

A possible solution for this issue is outlined below (see Figure 5). In our vision, information on target snow depth and respective indicators for blade (and potentially tiller) adjustment can be projected directly onto the snow surface in real time. This allows the operator to keep their eyes on the terrain in front of the vehicle.

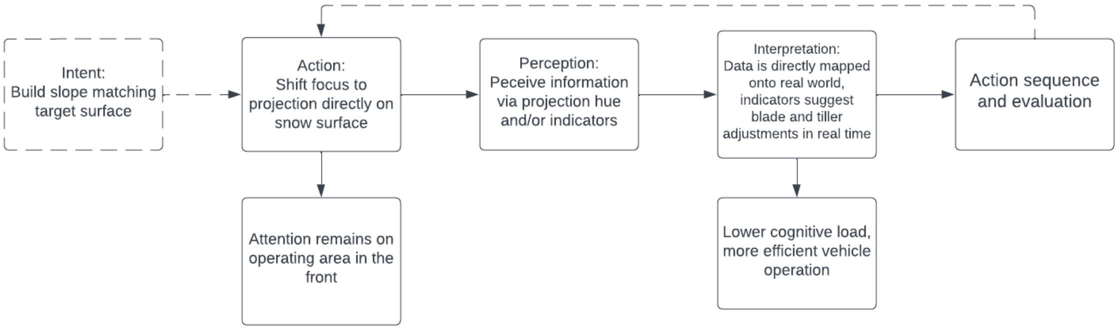

Figure 5: Conceptual interaction sequence model for UC1 scenario 1.

Scenario 2 presents handling with the very common threat or being involved in accidents, whether with obstacles (e.g., snow cannons or other equipment left behind on the slope) or skiers that regularly still use the slope even after closing hours. As visualized in this model (see Figure 6), to detect obstacles or skiers and react accordingly, operators need to check the front area in a long distance, mirrors, sides and, if available, the rear-view camera of their vehicle near constantly. This obviously demands high alertness from the operator and takes away a lot of attention from the main operating area in the front of the vehicle.

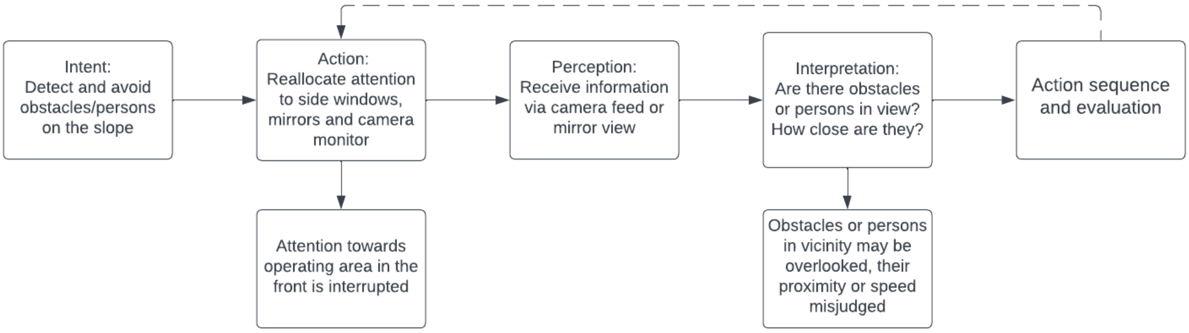

Figure 6: Initial interaction sequence model for UC1 scenario 2.

A possible solution for this is shown in Figure 7. When obstacles, e.g., skiers or animals in vicinity of the vehicle are detected, their presence could be visualized via ambient lights in the front of the cabin and/or light bars around the frame of existing onboard system screens. The type of obstacle (immobile object, animal, human, etc.) and its proximity could be conveyed via hue and size of the lit-up area (red and large area may indicate a human in proximity, yellow and small area may indicate an immobile object further away). The ideal visualization patterns for conveying this information still need to be investigated.

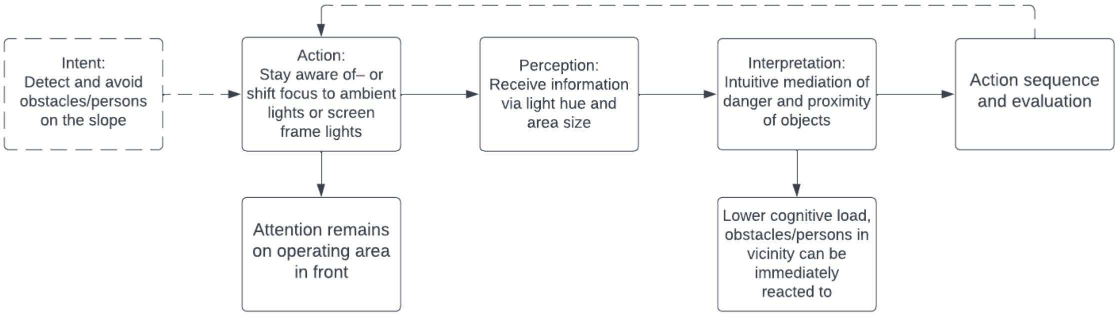

Figure 7: Conceptual interaction sequence model for UC1 scenario 2.

Scenario 3 (see Figure 8) presents a dangerous situation in which the operator is faced with fog, heavy snowfall or a snowstorm, which may significantly reduce vision or take it away entirely (whiteout). The operator may then try to navigate using known landmarks (e.g., trees, slope borders, immobile snow cannons) in combination with onboard map systems (if available) or, especially in a whiteout scenario, they must stop the machine and wait the situation out for their own safety. Navigating the slope without knowledge of location and/or orientation may result in major accidents, putting both vehicle and operator at risk.

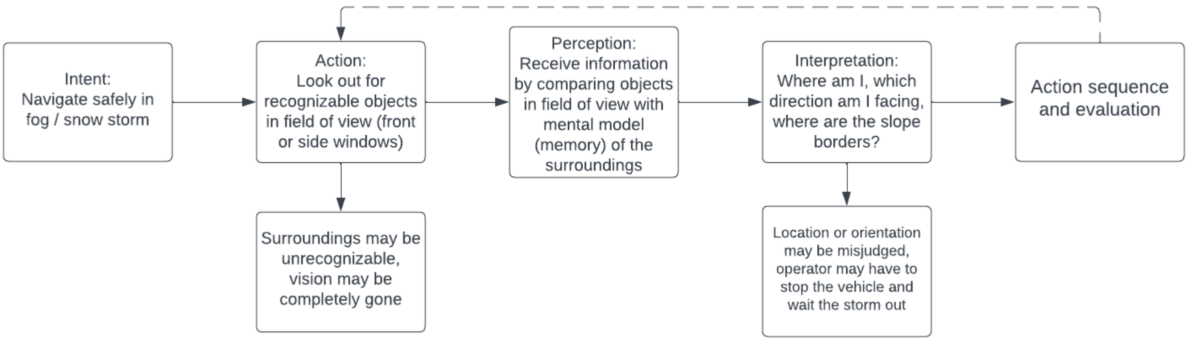

Figure 8: Initial interaction sequence model for UC1 scenario 3.

A possible solution for this scenario would be the projection of real-time map data, navigation and orientation information directly onto the windshield (or other surfaces in the cabin), onto the snow in front of the vehicle or even a projection into the fog ahead (see Figure 9). The operator could rely completely on the guidance provided by the onboard XR system to navigate to safety or continue their snow-grooming routine, even in a complete whiteout scenario.

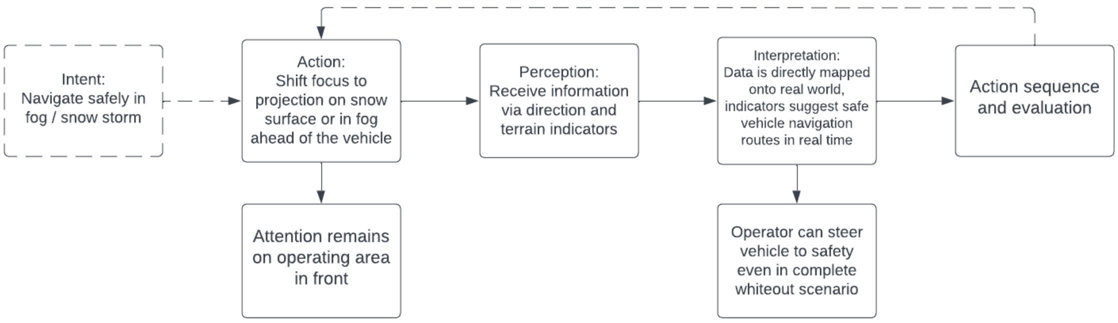

Figure 9: Conceptual interaction sequence model for UC1 scenario 3.

2.3 Use-case 2 interaction sequence models

In use case 2, logistics, we modelled the initial situations and conceptual interaction flows of the following scenarios:

1. Picking up container (from ground or stack)
2. Placing container (on ground or stack)
3. Collision avoidance

These scenarios were subject to user research in Hanko, Finland, in August 2023 and Ljungby, Sweden, in December 2023 within T3.1.

All three scenarios are embedded in the reachstacker operators' main workflow with only minor differences in execution depending on whether a container needs to be picked up from or placed on the ground, a stack or truck trailer respectively. The first scenario, picking up a container, requires the operator to approach the container with the vehicle, positioning the spreader accurately above it, locking the twist locks, raising the container and bringing it into a safe driving position for moving it to its target location. A centred frontal approach is ideal to make the spreader (container gripping attachment) positioning easier and more efficient, but depending on the location of the container, that may not always be possible. When positioning the spreader above the container, assessing the necessary movements and adjustments of vehicle, boom and spreader to align the spreader with the container precisely can prove difficult due to the reliance on visual estimation (see Figure 11). Once the spreader is in the right position, locking the twist locks happens automatically or with the press of a button. The status of the twist locks is visualized with coloured lights on an upper beam in the cabin and on the spreader itself (see Figure 10). While learning the basic mechanics of operating the vehicle and picking up containers is easy, doing this job efficiently requires a lot of experience.

Figure 10: Reachstacker cabin view, spreader visible in the center. Green lights indicate locked twist locks.

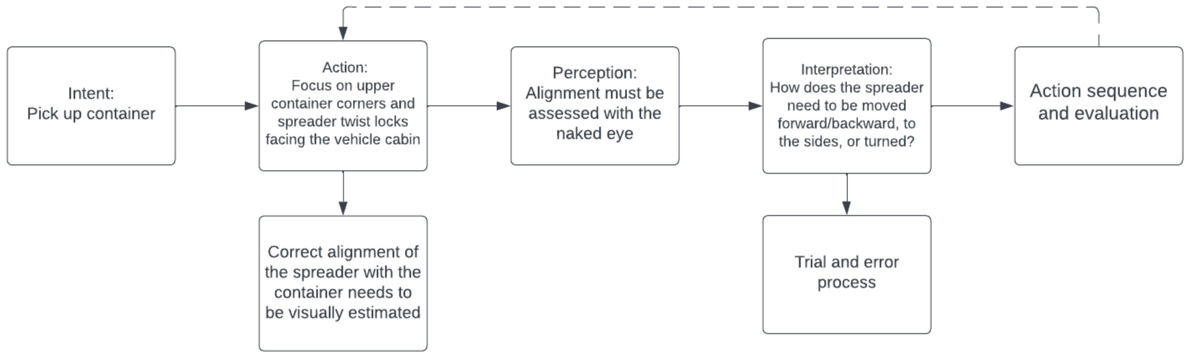

Figure 11: Initial interaction sequence model for UC2 scenario 1.

A possible solution for this would be a projection of transparent beams extending downwards from the twist locks in XR (via headset or augmented camera feed). Such a visualization could make the visual estimation of the correct positioning of the spreader above the container much easier and more efficient. The colour of these XR beams could indicate the correct positioning of the twist locks above the container corners, or their incorrect positioning respectively (see Figure 12).

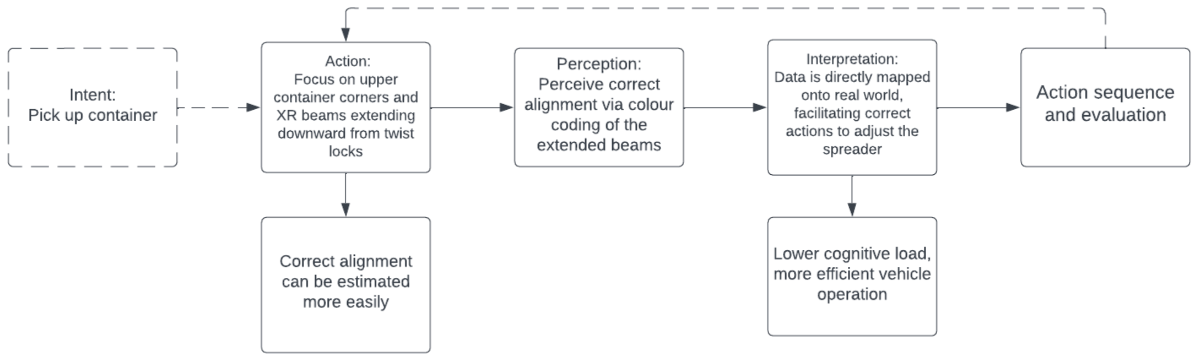

Figure 12: Conceptual interaction sequence model for UC2 scenario 1.

The second scenario is essentially a similar process as the picking up but focusing on the lower corners of the container instead (see Figure 13). When the container is positioned correctly, the twist locks need to be manually unlocked via pressing a button. Unlike the container pick-up process, where the twist locks would audibly “snap” into the openings in the corners of the container, there is no such additional feedback of a correct placement when placing a container.

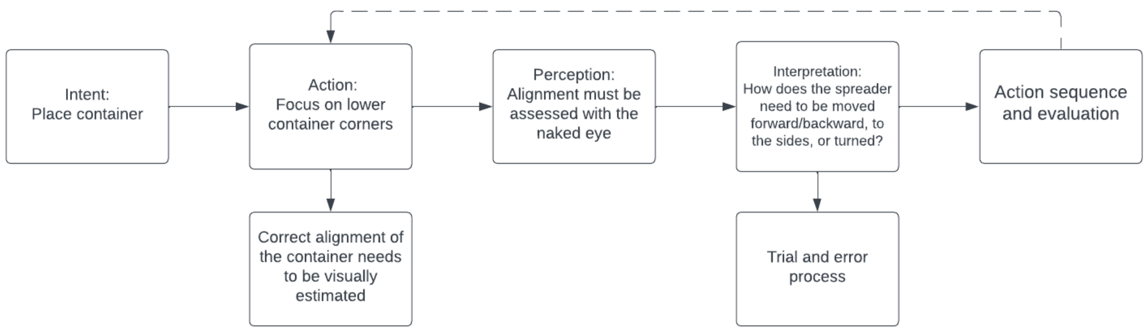

Figure 13: Initial interaction sequence model for UC2 scenario 2.

Here, the same solution as in scenario 1 would apply, with the simple adjustment of the XR beams extending downward through the picked-up container and into the ground (see Figure 14). This would facilitate the placement process just like the positioning of the spreader when picking up a container.

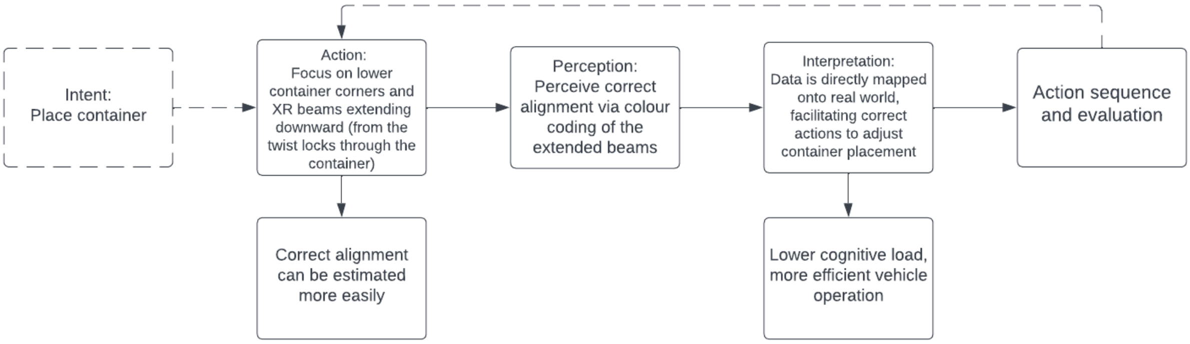

Figure 14: Conceptual interaction sequence model for UC2 scenario 2.

Finally, scenario 3, the checking for objects, persons and vehicles in the vicinity of the reachstacker happens repeatedly, usually after the completed picking-up and placement processes, whenever the reachstacker is being driven to another location. Checking the vicinity requires looking into the side mirrors (front or middle position, depending on the cabin position) and the rear window (see Figure 15). In some machines, rear cameras may be available, but we did not see an example for this on the user research visits. However, we expect this to be a similar technical setup as in snow-groomers or excavators (see Figure 6 and Figure 20).

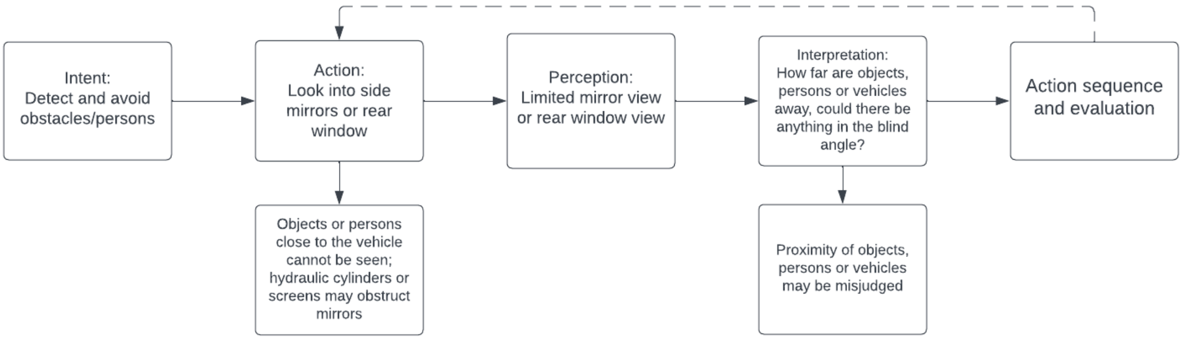

Figure 15: Initial interaction sequence model for UC3 scenario 3.

To solve this problem, objects such as other vehicles or persons moving around the container terminal could be detected via thermal cameras both from the vehicle or from large stationary floodlight posts overlooking the area. Detected objects could be identified and highlighted in XR view, e.g., by making their contours visible behind rows of container stacks or indicating their relative position to the vehicle with directional markers in the operator's field of view. This way, operators would need to check the mirrors much less often, perceiving other vehicles or persons in the area passively (see Figure 16).

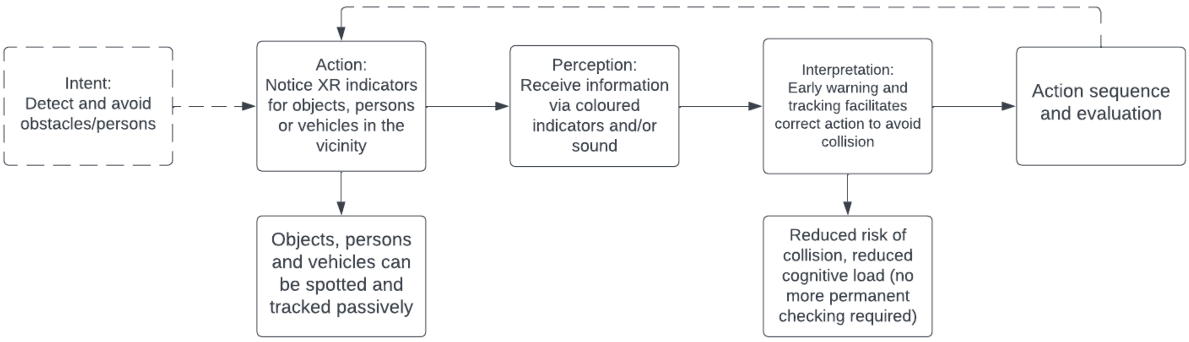

Figure 16: Conceptual interaction sequence model for UC2 scenario 3.

2.4 Use-case 3 interaction sequence models

In use case 3, construction, we modelled the initial situations and conceptual interaction flows of the following three scenarios:

1. Finish grading
2. Obstacle and person detection and avoidance
3. Infield design

These scenarios were subject to user research in Dresden and Glauchau, Germany, in May and August 2023 within Task T3.1.

The first scenario again represents the main job for the excavator operator: Creating an even surface at the target depth for an excavation pit or trench (see Figure 18). This involves excavating a previously defined pit or trench (see scenario 3) layer by layer without damaging any structures (e.g., power, gas or water lines) possibly hidden below that target depth. Creating an even surface at the exact target depth is a challenge that only very experienced operators will master without technological assistance systems or manual labour (finishing the surface by hand with a shovel). For less experienced operators or trainees, assistance systems are sometimes available, but these require near constant checking of the bucket's current depth in relation to the target depth, usually indicated on an onboard screen (see Figure 17). Paying attention to both screen information and real movement of the boom and bucket thus requires shifting the focus repeatedly, slowing down the operation significantly.

Figure 17: Excavator cabin view.

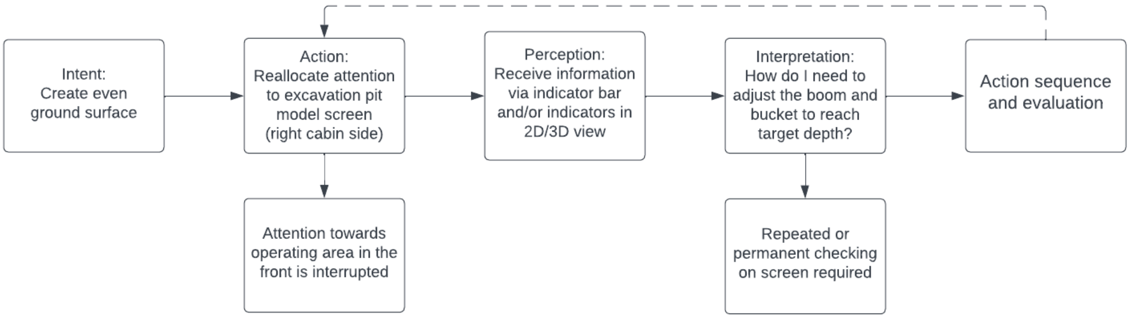

Figure 18: Initial interaction sequence model for UC3 scenario 1.

A possible solution for this would be using an LED panel on the boom or a laser projection on the ground to visualize the bucket's depth in relation to the target depth close to the area where the operator would usually focus (see Figure 19). This would remove the need to shift focus and make vehicle operation using such an assistance system much more efficient.

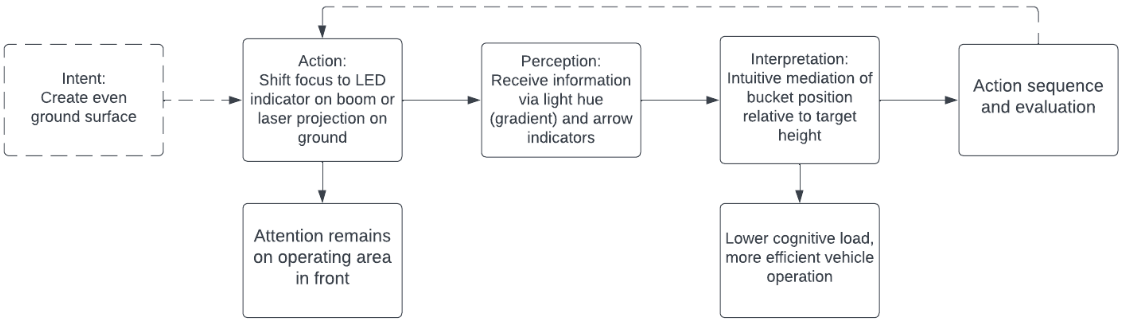

Figure 19: Conceptual interaction sequence model for UC3 scenario 1.

Scenario 2 resembles the obstacle and person detection and avoidance discussed in UC1 scenario 2. In this use case, other workers, vehicles and possibly pedestrians passing (or trespassing) the construction site need to be recognized and avoided (see Figure 20). Visibility is especially limited on the right side of the vehicle behind the boom.

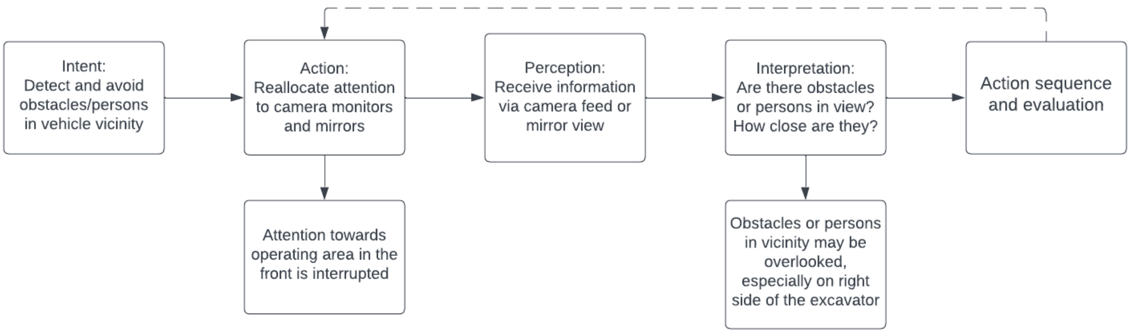

Figure 20: Initial interaction sequence model for UC3 scenario 2.

This scenario can be approached in the same way as UC1 scenario 2 using ambient lights around the cabin front and/or onboard system screens (see Figure 21).

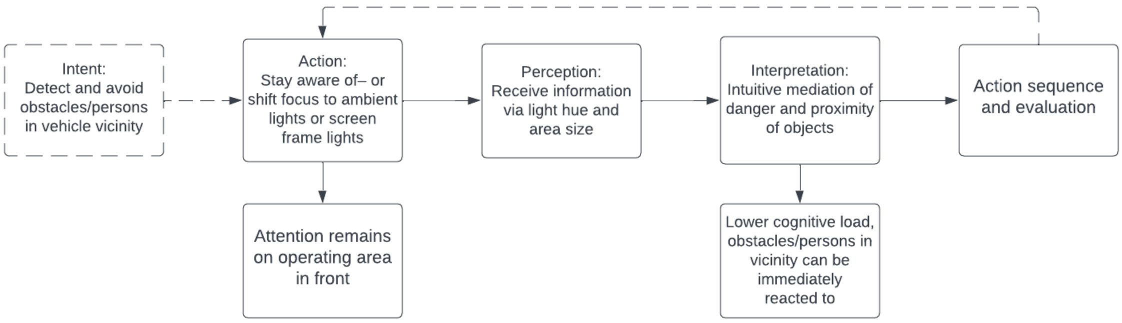

Figure 21: Conceptual interaction sequence model for UC3 scenario 2.

Finally, scenario 3 involves the creation of a digital excavation pit/trench model using the onboard systems, a task also called infield design (see Figure 22). This requires the corners of the planned excavation pit to be set up using real world markings (e.g., using rods, bricks, etc.) which must then be approached with the excavator bucket to calibrate the system and register the pit dimensions in the digital model. This is usually a complicated and time-consuming process that may require a complete restart if errors or inaccuracies are detected.

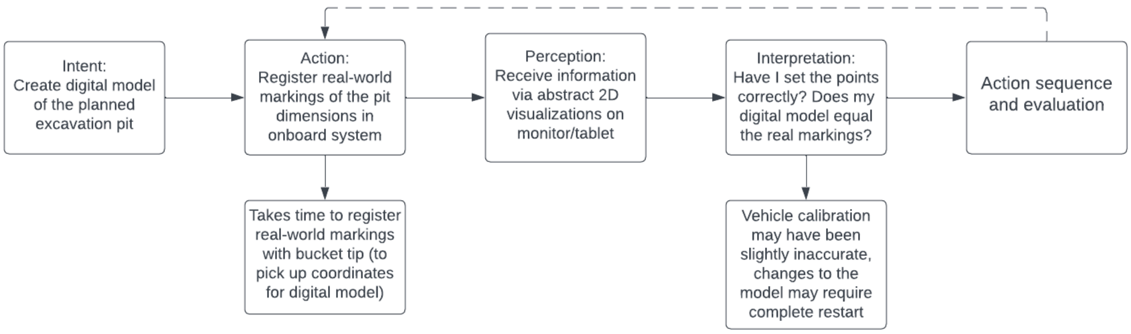

Figure 22: Initial interaction sequence model for UC3 scenario 3.

This can be solved by using LiDAR to register the corners of the planned excavation pit, projecting the outlines of the defined pit directly onto the ground or augmenting a camera view of the excavation pit with these outlines (see Figure 23).

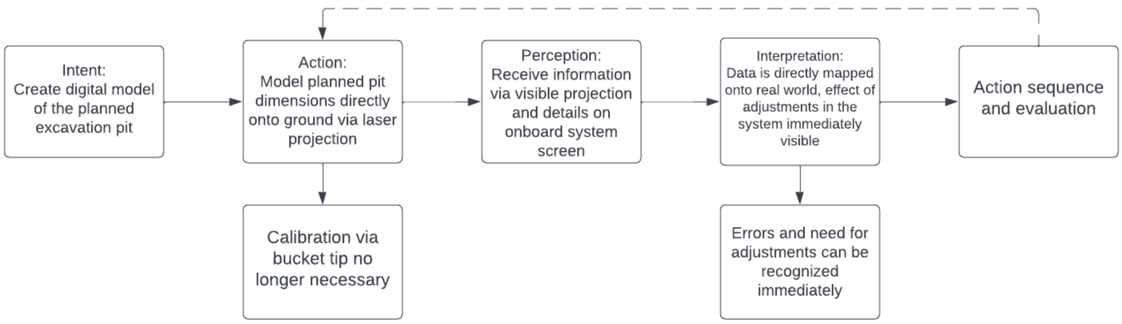

Figure 23: Conceptual interaction sequence model for UC3 scenario 3.

3 Concept design for XR interfaces

The focus of Task T4.3 is on the development of a concept design for future HMIs in off-highway machinery based on the previously identified sequence models. These sequence models will provide the developers a first draft overview of the possible requirements for these future HMIs. The following subsections will shortly describe how these sequence diagrams will form the basis for the different concepts of these to be identified and developed HMIs.

3.1 Multimodal interaction conceptualization, prototyping and evaluation

The interaction sequence modelling approach is part of the theoretic discussion of extended realities in the context of human—machine interaction. As this project aims at providing effective means of improved information presentation through XR technology, it must systematically explore, describe and evaluate different technologies and forms of information presentation.

We explore XR technology in our ideation activities and workshops that already spawned promising applications of XR interaction. However, XR interaction must be discussed in the context of human perception and processing capabilities and processes. The approach of interaction sequence models provides a framework that allows for a formalized description of XR interaction as part of actual machine operation tasks, including human states, goals and behaviour. Thus, this description is valuable in comparing different XR approaches and to highlight potential benefits in comparison to conventional information presentation or interaction techniques. We already started and will continue to use these modelling approaches to document our XR approaches and to work out their potential benefits for machine operation. Therefore, the interaction sequence models are considered a performant tool to connect our multi-disciplinary discussion that covers technological, contextual, human, and interaction related topics. It helps to reveal benefits of XR technology enabled interaction and to identify areas of improvement (Tasks T4.3, T4.4 and T4.5).

It supports the effective Scenario-based co-design of interaction concepts (Task T3.2) as it identifies touchpoints where our interaction concepts impact operators that can be used to deepen discussions and research in our co-design approaches.

As it implements specific ways of XR interaction in task contexts and goal-driven behaviour models it also highlights relations between specific XR interaction designs and operation performance. Relations, that should be considered in evaluating our XR approaches and the necessary means of prototyping (Task T3.4, Task T3.5.

This first approach on theory-informed interaction modelling is at first a tool to facilitate discussions among the different domains as it helps to discuss implications of XR technologies as XR interaction concepts in the context of real operation tasks and cognitive processes of operators. However, it lays the conceptual and theoretical fundament to further discuss multimodal and XR interaction in the context of important factors such as workload, ease-of-use and usability that are important criteria for successful integration of XR technology. We therefor continue our elaboration of this modelling approach to increase its descriptive capability, especially regarding contextual and situational factors that impact information perception, workload, and interaction effectiveness.

3.2 Multimodal remote operation XR interface

There are multiple ways to approach to exploit sequence model for the design of XR interfaces for the remote control of off-highway machinery, and they are deeply based on technological assumptions.

The first approach is to exploit an XR environment based on the digital twin of the area of operations. The main assumption in this case is that the vehicle is provided with Lidar sensors that can recreate an accurate and reliable virtual copy of the real environment. This Mixed-Reality concept is based on Varjo XR-3, which has high-quality pass-through cameras and allows the user to see the actual remote operation station together with the virtual representation of the reachstacker and harbour environment. The user controls the reachstacker with a Logitech G29 steering wheel and a Logitech Extreme 3D joystick and can reposition inside the virtual environment with hand gestures.

A second approach relies on a system of 360° cameras video streams, on top of which augmented content is rendered. The augmented reality content is based on several available sensors, such as the vehicle’s IMU system, and the twist locks pressure sensors. The system also makes use of other technologies, such as a computer vision model, to recognize objects in the video stream and show related information to the user. The user can choose between four different camera views: 1) cabin view/driver’s perspective, 2) spreader top-down view, 3) rear view/counterweight, and 4) drone/bird’s eye view. Visualization is done with an XR powerwall. This approach focuses on the pick-up and release of the containers, rather than driving the vehicle. The user controls the reachstacker’s boom and spreader with a Logitech Extreme 3D joystick. It is possible to rotate the camera view and change the video feed with hand gestures.

Evaluation of the described systems will be carried out during Task T3.4.

3.3 Haptic Interfaces

The stated objectives of proposing novel interaction sequence models relating to the in-vehicle HMIs in the THEIA^{XR} project are to:

1. Facilitate retrieval and interpretation of information relevant to the task. (e.g., to increase situational awareness by reducing distractions from the task at hand and from the external environment during information retrieval).
2. Facilitate planning and execution of a sequence of tasks aimed at achieving a formulated task goal.

In this regard, haptic feedback can take on two possible roles. The haptic feedback can either be *informative*, or it can be *assistive*.

3.3.1 Informative haptic feedback

In the case of informative haptic feedback, tactile [4] or kinesthetic [3] cues are conveyed to the operator with the aim of providing relevant information. In this case, the benefit of using haptics over e.g., vision or audio are three-fold:

First, certain information components are inherently easier to convey or to understand via the haptic modality (See "*modality adequacy*" in Figure 24). This is e.g., the case for external forces acting on a manipulated tool. Providing a replication of these forces at the handle of the input device used to control the tool provides a highly intuitive way of controlling tool-environment interaction forces [5].

Second, according to Wickens' multiple resource theory (MRT), a well-accepted model of perceptual and attentional resources in task performance, it is sensible to divide information feedback across multiple modalities to reduce cognitive load [6] (See "*perceptual bandwidth*" in Figure 24). Indeed, off-highway vehicle operators must already allocate a significant amount of their visual attention to the external environment around the vehicle, as well as a fair amount of auditory attention to the sounds of the vehicle itself and sounds from the surroundings to ensure safe and effective operations. Thus, providing additional information feedback via the visual and auditory channels runs the risk of overloading the operator's sensory processing capabilities for these modalities, resulting in degraded task performance. There may thus be a larger margin to be exploited in operators' haptic information processing capabilities. This could allow for an overall larger volume of information to be transmitted to and processed by the operator, without significantly increasing cognitive load, and without risking degrading performance [5].

Finally, haptic cues are inherently private [7]. This means that information destined only for the operator can safely be provided through this modality.

3.3.2 Assistive haptic feedback

The other role that haptics can take on in an improved in-vehicle HMI is to actively assist the operator in performing a task (see "*Assistive haptic feedback*" in Figure 24). This approach is commonly referred to as "*virtual fixtures*" [8] when it acts as a restriction to operator motions, or "*haptic shared control*" ([9], [10]) when an automatic controller steers the operator's motions towards a target. This type of feedback is necessarily kinesthetic in nature, as it implies that operator motions must be actively guided in some way.
The role of haptic feedback as active assistance has been extensively explored over a wide variety of domains such as automotive [10], aerospace [11], medical 12/22/2023 9:31:00 AM or micromanipulation ([12], [13]), to cite a few. It should be noted that assistive haptic feedback also necessarily is informative, as the operators learn to associate the perceived assistive forces and torques with the state of the manipulated tool, the state of the external environment, or the state of the task at hand (dotted red arrow in Figure 24). Similar improvements in cognitive load as those expected from purely informative feedback are also observed during the use of assistive haptic feedback [10].

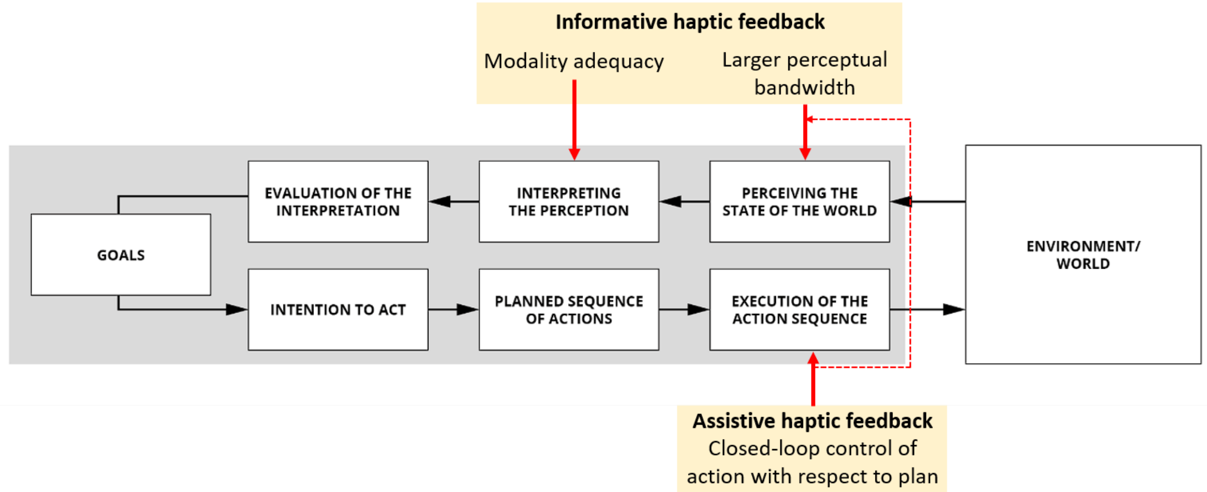

Figure 24: Levels of haptic feedback components in the basic interaction sequence model

Haptic feedback components can act at three levels of the basic interaction sequence model. First, informative haptic feedback can provide a better or greater perception of the environment, leveraging the principles of MRT to give the operator a larger perceptual bandwidth. Second, informative haptic feedback can ease interpretation of environment perception by providing a feedback modality that is more adequate to the information being conveyed. Finally, assistive haptic feedback can actively improve the execution of a desired action sequence, by actively guiding the operator.

3.3.3 Haptic feedback possibilities in UC1

Build slope matching target surface

When considering the role of haptics, we can classify the tasks described in the UC1 interaction sequence model (Sec. 2.2) into two main groups: “*blade manipulation*” (« preparing a smooth surface », « adding and removing snow », « moving excess snow to other parts of the slope » and « avoiding damages to the surface »), and “*control of vehicle trajectory*” (« maintaining slope borders »). Both these task groups imply perception and action, hence both informative and assistive haptics can be envisioned for them.

In the case of blade manipulation, simple informative feedback could be provided in the form of notifications signaling that the blade has reached the ideal pose or that it is travelling towards a problematic pose. This may be realized either through tactile feedback (e.g., vibration of the primary input device, which should be designed considering internal vehicle vibrations during operation [14]) or kinesthetic feedback (change in input device apparent stiffness or damping [15]). More complex feedback could take the form of guidance information, either through multipoint tactile stimulation capable of encoding direction and distance information [16], directional skin-stretch tactile feedback [17], or kinesthetic feedback [9]. In the latter case, the feedback would also become assistive in nature. Since the snow groomer’s blade moves with 4 + 2 degrees of freedom (three rotations of the main blade, one translation of the main blade, and two independent opening/closing rotations of the blade wings), any implementation of haptic guidance should provide a natural and intuitive mapping to these degrees of freedom.

An alternate or complementary way in which haptics could enhance the perception of relevant information relating to the blade’s interaction with the snow would be to reflect forces acting on the blade onto the handle, i.e., implement conventional force feedback teleoperation of the blade [12]. While conceptually rather simple, such an approach poses two major challenges: that of directly measuring the forces acting on the blade (integration of multiple force sensors) or of inferring it from other data (which requires additional research), and that of ensuring stability of the force-feedback scheme.

A final interesting concept would be haptic noise cancellation [18]. Since snow groomer operators attempt to finely control the blade position in an environment subject to shocks and vibration, actuating the input device handle in such a way as to cancel out as much of this mechanical noise as possible could provide the operator with a clearer feeling of how they are actually controlling the blade. However, this remains highly hypothetical, and the feasibility of this approach largely depends on the bandwidth and amplitude of this mechanical noise as well as on the achievable dynamics of the novel haptic input device designed within the project.

Collision avoidance

As described in UC1 (see Use case 1 interaction sequence models), collision avoidance comprises two sub-tasks, "*obstacle detection and recognition*" and the "*obstacle avoidance*" per se.

The "*obstacle detection/recognition*" is a purely perceptual sub-task. Hence, only informative haptic feedback makes sense in this context. Again, the most basic form of haptics that can be imagined assisting users in perceiving the presence of an obstacle is notification feedback, e.g., through simple vibration of an element in contact with the user. Tactile notifications have been shown to improve reaction times in certain safety-critical contexts [19]. In addition to notifying the operator about the presence of an obstacle, more complex types of feedback could be imagined to haptically describe the type of obstacle, e.g., through a specific set of tactile patterns [20]. Similarly, tactile patterns could encode the distance and direction of the obstacle.

The "*obstacle avoidance*" sub-task is a vehicle control task, hence both informative and assistive haptics are sensible. Indeed, we could envision only providing real-time feedback of their distance and heading relative to the vehicle as previously mentioned, which would in theory be sufficient for an operator to avoid the obstacle. However, more active forms of assistance could be imagined either in the form of virtual fixtures keeping the operators away from input commands that would lead to a collision, or haptic shared control guiding the operator towards a computed evasion course. This last solution, however, would imply that in addition to detecting obstacles, a path planning step would have to be executed.

Zero-sight navigation

Zero-sight navigation is a vehicle control task which can be considered a generalization of the "obstacle avoidance" sub-task discussed above and therefore does not require specific discussion.

3.3.4 Haptic feedback possibilities in UC2

While neither the context of use and work environment threat analysis described in Deliverable 3.1, nor the interaction sequence models for UC2 in the present document are available, we can still hypothesize that feedback approaches proposed in UC1 may also be useful here. This is particularly the case for collision and obstacle avoidance, as well as the ideas of feeding back interaction forces between tool (container or gripper) and environment at the input device handle.

Interaction forces when manipulating a container

If possible, feeding back interaction forces when manipulating a container could provide reachstacker operators with a more intuitive understanding of load distribution within the container, possible imbalances, as well as serve the purpose of providing naturalistic impact cues which operators could use to finely place containers or avoid excessive damage due to collisions. However, the same challenges as discussed above apply in this scenario.

Obstacle and person detection and avoidance during container manipulation

This is conceptually identical to the "*obstacle avoidance*" scenario in UC1. However, some particularities of the reachstacker use-case need to be considered. Indeed, while obstacle avoidance in UC1 hinges on steering the entire vehicle away from any obstacle, the reachstacker can extend its arm and gripper, pick up or release containers and possibly rotate them. This behavior effectively changes the geometry of the vehicle depending on the stage of the task being performed. Beyond simple steering of the vehicle to avoid obstacles, it is thus necessary to consider potential collisions that may occur because of the motions of the arm, gripper and manipulated container.

This doesn't profoundly impact notification feedback, although we could envision different haptic notification systems or signals dedicated to warning about a vehicle-environment collision, an arm/gripper-environment collision, or a container-environment collision. However, it means that guidance feedback ideally needs to be provided at the level of the interface currently in use for controlling the arm and gripper.

3.3.5 Haptic feedback possibilities in UC3

Although Haption is not actively involved in developing haptic feedback solutions for UC3, it is worth highlighting some conceptual similarities with goals and tasks in other use-cases which may indicate a potential for benefits from haptics in this use-case.

Finish grading

In the case of finish grading, the goal of the task is to guide the excavator bucket along a defined trajectory. Thus, insights gained from the "*building a slope matching target surface*" scenario of UC1 may very well find applications here in the future.

Obstacle and person detection and avoidance

This task is conceptually identical to the "*obstacle avoidance*" scenarios in UC1 and UC2 could thus benefit from the same forms of haptic feedback. However, similarly to the discussion presented above for UC2, an excavator is also a vehicle which can actively change its overall shape during operation and could require separate consideration of vehicle-obstacle collision and boom-, arm-, or bucket-obstacle collisions.

Infield design

In the future, this task may benefit from haptic notifications of successful or unsuccessful actions during the registration phase. This could help in making the interaction more seamless. In this case we could borrow from the concepts for tactile notification discussed above. It should be noted that even if the interactions required for this registration phase are with another HMI component than the primary input device, it may still be sensible to provide feedback at the primary input device as even distal feedback has been shown to yield better interaction performances (e.g., [21]). The infield design task also could benefit from similar haptic guidance as the finish grading task.

3.4 Visualization Infrastructure and Interaction System Interface

The sequence diagram representations provide us an insight in the XR interfaces and technologies that will be targeted within the project. Therefore, it is important to know the interfaces from the different technologies and how at the end, they need to be integrated in the cabins of the mobile machinery. Additionally, all these technologies will require enough communication and data interfaces to interact with the machine and the environment, to provide the data to the human operator in the cabin and to be able to process the data directly on the machine, potentially without being dependent from an external computing solution (e.g., cloud).

The visualization infrastructure and interaction system development will focus mainly on bringing the technologies into the machines and directly into the cabin where the human operator is controlling the machine. Therefore, it is important for this part to receive the information from the sequence models which interaction patterns and methods are being used and what kind of technology (i.e., hardware and/or software) will be required to inform the human operator and to make the non-perceivable perceivable. The sequence models will provide first information about the information required for the human operator, which will in turn result in the technologies needed to actually provide this information. Therefore, this development will follow closely the results coming from these models.

Continuing on this, the technology providers/developers will closely interact with the visualization infrastructure and interaction system provider. First interface design steps have already been done, to identify what is needed to make it possible to finally integrate the technologies into the mobile machinery and provide the information to the human operator while operating this machinery.

Finally, the implementation of the visualization infrastructure will take place (mainly in Task T5.3 of the project). This will take all the information as basis and will provide an infrastructure capable of hosting the XR applications based on the requirements of the operators and additionally with enough interfaces to interact with the various hardware (e.g., sensors and actuators) to provide the required information coming from the sequence models and the questionnaires performed with the human operators. And concluding, the infrastructure will then be included in the actual demonstrators, which are the mobile machinery.

This work in this task will form the basis for the technical development taking place in Task T5.3 and final integration in Task T6.1.

4 Conclusions

The interaction sequence models show current and possible future information retrieval cycles for the main use case scenarios. The initial interaction sequence models visualize how in many of these use case scenarios the act of looking at a screen interrupts the operators' focus on the main operating area, especially where precise adaptation to a continuously changing situation is needed. The conceptual interaction sequence models show possible technical solutions that are being explored in the project. They also show how in some cases, the same technical developments can be used in different use-cases: Ambient lights may highlight obstacles and persons in the vicinity or vehicles, laser projections on the ground may be used to light the way or indicate tool adjustments. Thus, we expect XR and information design developments made in this project to be generally applicable to other off-highway domains as well, such as agriculture or mining.

The XR interaction concepts described by these sequence models provide a basis for technological discussion, exploration and development of solutions, as the last chapter shows. Prototyping and/or development of some of these solutions has already started. In the next months, along with the progress of WP3 and WP4 in general, the interaction sequence models will be updated and elaborated.

References

[1] D. A. Norman, 'Stages and levels in human-machine interaction', *Int. J. Man-Mach. Stud.*, vol. 21, no. 4, pp. 365–375, Oct. 1984, doi: 10.1016/S0020-7373(84)80054-1.

[2] D. Norman, 'Cognitive Engineering', in *User Centered System Design: New Perspectives on Human-Computer Interaction*, 1986, pp. 31–61. doi: 10.1201/b15703-3.

[3] D. Szafir, 'Mediating Human-Robot Interactions with Virtual, Augmented, and Mixed Reality', in *Virtual, Augmented and Mixed Reality. Applications and Case Studies*, vol. 11575, J. Y. C. Chen and G. Fragomeni, Eds., in Lecture Notes in Computer Science, vol. 11575. , Cham: Springer International Publishing, 2019, pp. 124–149. doi: 10.1007/978-3-030-21565-1_9.

[4] L. Jones and N. Sarter, 'Tactile Displays: Guidance for Their Design and Application', *Hum. Factors*, vol. 50, pp. 90–111, Mar. 2008, doi: 10.1518/001872008X250638.

[5] B. Weber and C. Eichberger, 'The Benefits of Haptic Feedback in Telesurgery and Other Teleoperation Systems: A Meta-Analysis', in *Universal Access in Human-Computer Interaction. Access to Learning, Health and Well-Being*, vol. 9177, M. Antona and C. Stephanidis, Eds., in Lecture Notes in Computer Science, vol. 9177. , Cham: Springer International Publishing, 2015, pp. 394–405. doi: 10.1007/978-3-319-20684-4_39.

[6] C. Wickens, 'Multiple resources and performance prediction', *Theor. Issues Ergon. Sci.*, vol. 3, pp. 159–177, Jan. 2002, doi: 10.1080/14639220210123806.

[7] D. Tam, K. E. MacLean, J. McGrenere, and K. J. Kuchenbecker, 'The design and field observation of a haptic notification system for timing awareness during oral presentations', in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Paris France: ACM, Apr. 2013, pp. 1689–1698. doi: 10.1145/2470654.2466223.

[8] L. B. Rosenberg, 'Virtual fixtures: Perceptual tools for telerobotic manipulation', *Proc. IEEE Virtual Real. Annu. Int. Symp.*, pp. 76–82, 1993, doi: 10.1109/VRAIS.1993.380795.

[9] M. O'Malley, A. Gupta, M. Gen, and Y. Li, 'Shared Control in Haptic Systems for Performance Enhancement and Training', *J. Dyn. Syst. Meas. Control-Trans. Asme - J DYN SYST MEAS CONTR*, vol. 128, Mar. 2006, doi: 10.1115/1.2168160.

[10] M. Steele and B. Gillespie, 'Shared Control between Human and Machine: Using a Haptic Steering Wheel to Aid in Land Vehicle Guidance', *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 45, pp. 1671–1675, Oct. 2001, doi: 10.1177/154193120104502323.

[11] D. P. Haanpaa and G. P. Boston, 'An advanced haptic system for improving man-machine interfaces', *Comput. Graph.*, vol. 21, no. 4, pp. 443–449, Jul. 1997, doi: 10.1016/S0097-8493(97)00017-4.

[12] C. Basdogan, A. Kiraz, I. Bukusoglu, A. Varol, and S. Doğanay, 'Haptic guidance for improved task performance in steering microparticles with optical tweezers', *Opt. Express*, vol. 15, pp. 11616–21, Oct. 2007, doi: 10.1364/OE.15.011616.

[13] A. Bolopion and S. Régnier, 'A Review of Haptic Feedback Teleoperation Systems for Micromanipulation and Microassembly', *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 3, pp. 496–502, Jul. 2013, doi: 10.1109/TASE.2013.2245122.

[14] J. Ryu, J. Chun, G. Park, S. Choi, and S. Han, 'Vibrotactile Feedback for Information Delivery in the Vehicle', *IEEE T Haptics*, vol. 3, pp. 138–149, Apr. 2010, doi: 10.1109/TOH.2010.1.

[15] D. Abbink and M. Mulder, 'Exploring the Dimensions of Haptic Feedback Support in Manual Control', *J. Comput. Inf. Sci. Eng.*, vol. 9, pp. 1–9, Mar. 2009, doi: 10.1115/1.3072902.

[16] J. B. F. V. Erp, H. A. H. C. V. Veen, C. Jansen, and T. Dobbins, 'Waypoint navigation with a vibrotactile waist belt', *ACM Trans. Appl. Percept.*, vol. 2, no. 2, pp. 106–117, Apr. 2005, doi: 10.1145/1060581.1060585.

[17] S. Norman, A. Doxon, B. Gleeson, and W. Provancher, 'Planar Hand Motion Guidance Using Fingertip Skin-Stretch Feedback', *IEEE Trans. Haptics*, vol. 7, pp. 121–130, Apr. 2014, doi: 10.1109/TOH.2013.2296306.

[18] S. G. Yuen, K.-A. Dubec, and R. D. Howe, 'Haptic noise cancellation: Restoring force perception in robotically-assisted beating heart surgery', *2010 IEEE Haptics Symp.*, pp. 387–392, Mar. 2010, doi: 10.1109/HAPTIC.2010.5444626.

[19] R. Gray, C. Ho, and C. Spence, 'A Comparison of Different Informative Vibrotactile Forward Collision Warnings: Does the Warning Need to Be Linked to the Collision Event?', *PLOS ONE*, vol. 9, no. 1, p. e87070, Jan. 2014, doi: 10.1371/journal.pone.0087070.

[20] S. Brewster and L. M. Brown, ‘Tactons: structured tactile messages for non-visual information display’, in *Proceedings of the fifth conference on Australasian user interface - Volume 28*, in AUIC ’04. AUS: Australian Computer Society, Inc., Jan. 2004, pp. 15–23.
[21] J. Henderson, J. Avery, L. Grisoni, and E. Lank, *Leveraging Distal Vibrotactile Feedback for Target Acquisition*. 2019, p. 11. doi: 10.1145/3290605.3300715.

ABBREVIATIONS / ACRONYMS

CPS	Cyber-Physical System
HMI	Human-Machine Interface
LiDAR	Light Imaging, Detection and Ranging
IMU	Inertial Measurement Unit
MRT	Multiple Resource Theory