Virtual Environment-Based Teleoperation of Forestry Machines: Designing Future Interaction Methods

Simon Westerberg
Umeå University, Sweden
and
Anton Shiriaev
Norwegian University of Science and Technology, Norway

Virtual environment-assisted teleoperation has great potential as a human-robot interaction paradigm for field robotic systems, in particular when combined with elements of automation. Unstructured outdoor environments present a complex problem with many challenging elements. For the specific application of forestry machines, we investigate which steps are required in order to implement such a system, what potential benefits there are, and how individual components can be adapted to efficiently assist forestry machine operators in their daily work in the near future. An experimental prototype of a teleoperation system with virtual environment-based feedback is constructed using a scenario-based design process. The feasibility of the implementation is partly verified through experimental studies.

Keywords: Human-robot interaction, teleoperation, virtual environments, field robotics, system design

1. Introduction

We study here a new interaction method for assisting the operator of a forestry machine. It is based on combining a specially developed visualizing virtual environment, telerobotics, and a developed control system that allows us to implement autonomous motions. We concentrate on various conceptual scenarios of using the interface in the nearest future as well as in the far future.

1.1 Motivation

The predominant method for forest harvesting in Northern Europe is cut-to-length logging, where the trees are handled by two machines: a harvester, which fells and delims the trees and cuts the trunk into logs, and a forwarder, which collects the logs in a load bunk, and carries them to the nearest road for transportation. Nowadays, these forestry vehicles are technologically advanced machines. However, currently, to achieve efficient usage, an operator requires good skills and long training that motivates further development of the machines. The difficulty is related to how the cranes of current forestry vehicles are controlled by the operator. The cranes are redundant manipulators with the position of the end effector (grapple or harvester head) determined by four variables...
Westerberg and Shiriaev, Virtual Environment-Based Teleoperation of Forestry Machines

(three angles of rotation and one linear displacement position), as shown in Figure 1a. Traditionally, the hydraulic cylinders associated with each of these variables are individually controlled by some motion of one of two joysticks. The typical mapping for a forwarder is shown in Figure 1b. For a harvester, a number of additional buttons allow to control different functions of the end effector.

![Figure 1a](a) A forwarder with joint variables.  
![Figure 1b](b) Current manual joint control with dual joysticks.

*Figure 1.* Typical mapping between joystick control and individual link movement of a forwarder crane.

One of the most severe problems related to crane control is pain and fatigue, mainly in arms, neck, and shoulders (Winkel, Attebrant, & Wikström, 1998). The current method of manipulating the crane requires that the operator almost continuously performs fine motor motions in a rapid pace. The high duration of these repetitive motions with low variation and not many natural breaks can be particularly harmful, since the same muscle fibers are used for long periods of time without enough time to recover.

The workload is not only physical, but the mental strain and stress is also substantial (Inoue, 1996). A large number of decisions are constantly performed and the complexity of the manipulator control may lead to mental capacity being used to perform low-level control of the individual links instead of higher-level planning, such as selecting which trees to cut or pick up. Apart from reducing performance, mental fatigue also increases the risk of accidents (Lilley, Feyer, Kirk, & Gander, 2002).

Whole body vibrations are also considered an important issue in the working environment for the operators (Jönsson & Löfroth, 2007). These vibrations are caused by the off-road driving and can cause physical consequences, e.g. lower back pain. The environment also contains a considerate level of noise in a large range of frequencies that can contribute to the mental stress and affect the operators’ performance (Winkel et al., 1998).

The individual joint control does not only affect the working conditions for the operator. It also increases the amount of time that is needed for an operator to learn how to maneuver the crane in a reliable manner, let alone to become a proficient driver able to determine to which extent the different actuators should be used in different situations and which motions are most efficient for each task. The efficiency by which the tasks are performed are, of course, affecting productivity and operating cost, and is thereby of great concern for the forestry machine owners. Another similarly important issue is the fuel consumption; a reduced fuel usage has both economic and environmental incentives. It is clear that improved human-machine interaction and a better control system are necessary to resolve these issues.
1.2 Related Work

Autonomous vehicles are suggested as a long-term goal for the forestry industry as a means to deal with the problems described above and to increase productivity (Hallonborg, 2003). Even though autonomous vehicles have been successfully implemented in other areas, such as mining (Bellamy & Pravica, 2011), cargo handling (Bae, Choe, Park, & Ryu, 2011), or agriculture (Shearer, Pitla, & Luck, 2010), fully autonomous forestry machines are very complicated to achieve for several reasons. First, they operate in unknown and unstructured environments under rough conditions, where sensors and other equipment must work reliably even when exposed to heat, cold, snow, and dirt. Second, a highly capable artificial intelligence would be required for making all decisions involved in the work task and even more so for appropriate responses to unexpected events. Furthermore, automation of all aspects of end effector control is not realistic with current crane designs, due to the passive degrees of freedom of the end effector, particularly noticeable when excited under large and heavy load. Further, it is not enough for an autonomous system to be merely functional, it would also need to be able to meet the currently high productivity demands.

Efforts have been made to automate one part of the work task, that is, transporting the timber from the felling site to the collection site (Hellström, Lärkeryd, Nordfjell, & Ringdahl, 2009; Ringdahl et al., 2011). By having an operator demonstrate the preferred path, the machine can repeat it and correct for obstacles. However, the majority of working time in a forwarder is spent on crane control. Since this also is a major cause of some working environment problems, it is interesting to see how automation can be utilized in this area.

One level of automation that can be realizable in a shorter term is to allow autonomous execution of certain simple crane motions, such as the repetitive loading and unloading motions (end effector work excluded). Automated motions not only relax the workload of the human operator and introduce natural breaks with time for muscle recovery, they can also be planned with the intention of optimizing some performance index. By incorporating the motion planning method described in (Mettin, La Hera, et al., 2009), it is possible to calculate a time-suboptimal trajectory for the crane from a given initial position to a given target point. Furthermore, Mettin, Westerberg, Shiriaev, and La Hera (2009) show that this method has the potential to improve productivity when compared to the standard manual control. If the cost function for the optimization is appropriately chosen, the resulting trajectory could be the one that reaches the target point with least fuel consumption.

Another approach for solving some of the aforementioned problems is teleoperation, defined as technology that supports physical action at a distance (Hannaford, 2000). This concept is used wherever it is considered useful to combine the computational efficiency, precision, or sturdiness of machines with the perception ability and decision-making processes of humans. As a result, teleoperation technology for robotic systems has been considered, developed, or used for a large spectrum of tasks, e.g., for work in mining (Duff et al., 2009), nuclear facilities (Qian, Song, Bao, & Zhang, 2012), agriculture (Murakami et al., 2008), surgery (Tobergte, Konietschke, & Hirzinger, 2009), space robotics (Yoon et al., 2004), and even surgery in space (Haidegger & Benyo, 2008).

Moving the forestry machine operator to a separate remote location could have several advantages. One immediate effect would be an improvement of the working conditions. The operator is no longer exposed to the noise and vibrations inevitably apparent in the cabin. The operator safety is also increased; by being located at a distance from the vehicle, the operator is less exposed to dangerous events like an overturning vehicle or a collision with a falling tree. Furthermore, the cabin is no longer needed, allowing the machines to be made lighter and smaller, which in turn could result in less fuel consumption and less negative effects on the environment. The vehicles can also be redesigned for maximal productivity and minimal building cost when operator safety constructions are irrelevant.

Teleoperation is not a new concept to the forestry sector. The Besten system (Bergkvist, Nordén,
& Lundström, 2006) is such an experimental prototype, where a number of forwarders take turns to control a cabinless harvester remotely. During periods of teleoperation, the forwarder operator is located at a short distance from the harvester and comprehends the environment by direct sight. This particular system is not the solution to all the aforementioned problems. Productivity analysis results show that the competitiveness of this multi-machine system is limited (Ringdahl, Hellström, & Lindroos, 2012). Further, since the operator uses the same manual crane and end effector control as a conventional machine, small gains are made in regards to the working environment. It does, however, demonstrate that removing the cabin allows for interesting possibilities of redesigning the machine and investigating new ways of optimizing the working routines. Teleoperation combined with semi-autonomous behavior, as well as with an appropriate interaction method, could further improve the physical and mental situation for the driver.

1.3 Classification of Teleoperation Systems

The wide range of application areas for teleoperation systems naturally imply a large variation of the conditions and situations in which they are used. Differences appear, for example, in the extent to which the environment/workspace is known in advance (and thereby can be used for planning), whether the environment is dynamic or static, to which degree the tasks are known in advance, and whether the priority lies in speed/operational time for the task or in safety/precision. This variety of conditions introduces different requirements and limitations on the teleoperation systems, which in turn has resulted in a multitude of different conceptual and technological solutions.

One way of classifying teleoperation systems is by the control method. Early teleoperation systems consisted of two identical or similar robotic manipulators where the angles of the individual joints of a master robot were directly used to control each joint of a slave robot (Kheddar, Chellali, & Coiffet, 2002; Hannaford, 2000). This direct control, or joint control, is a simple method that can even be conveyed without electric control by using only mechanical connections between the manipulators. For a robot with a relatively simple kinematic design, this method may turn out successful. However, for robots and manipulators with complex kinematics, it can be advantageous to avoid transferring the complexity to the user control interface. In many cases, the individual joint configurations are not of great importance and can be automatically handled by some inverse kinematics scheme, letting the user focus on the position of the end effector, e.g., guiding it along a desired path. In addition, this does not confine the user interface to a kinematic copy of the remote manipulator. This simplified computer-assisted manual control method is called Cartesian control or end-effector control. In other situations, even the end effector path is of subordinate importance or is for some other reason best handled by a computer. This leaves the user with the task of planning, describing, and communicating high-level commands, or tasks, to the remote robot. This is known as supervisory control (Sheridan, 1992).

Teleoperation systems also have different methods for providing feedback to the user. Some teleoperation systems use streaming video as feedback, which can be considered a cheap and easy solution. Another option is to use a virtual environment (VE) to present a virtual dynamic model of the remote location. This method has several advantages (Kheddar, Neo, Tadakuma, & Yokoi, 2007), considering that a VE can provide:

- An image quality and update rate that is not dependent on the environment sensors/cameras.
- An arbitrary point-of-view. Since objects are modeled, they can be projected onto any camera view; the view point is not fixed to wherever a camera can be mounted.
- A more task-oriented and intuitive user interface, e.g., the user may instruct a robot to pick up an object by selecting the object in the VE.
- Visualization of the intentions of the system in automated tasks, e.g., objects that a robot intends to pick up can be highlighted in the VE.
• System and operator share a “common language.” Actions can be linked to objects instead of pixels or coordinates.
• Lower bandwidth requirements by communicating abstractions (objects and their properties) instead of large video data streams.

There is, however, a clear limitation to when VE feedback can be used, namely that the system must include a model of the environment. This either means that the teleoperation must be performed in known environments or that the system is equipped with sensors and an environment recognition system capable of interpreting sensor data for modeling the environment. Recognition and classification of an unstructured outdoor environment is still an open problem, although a number of methods solving particular subclasses of this problem have been presented, e.g., the detection of trees (Forsman & Halme, 2005; Kulovesi, 2009) and logs (Park, Shiriaev, Westerberg, & Lee, 2011). There are also methods for localization in unknown environments, e.g., (Miettinen, Ohman, Visala, & Forsman, 2007), as well as for terrain classification (Munoz, V, & Hebert, 2009).

1.4 Aim of the Paper

The long-term goal of VE-based teleoperation and semi-automation of forestry machines is based on the hypothesis that it allows the productivity to be maintained or increased while the workload for the operator is reduced. This paper aims to help validate this claim by specifically evaluating the effects of VE feedback on different aspects of teleoperation performance. Further, we aim to investigate what components a successful VE teleoperation system would require. We want to know which incremental steps can and should be taken to reach the goal and in which order they can be applied to be as useful and beneficial as possible along the way. For the design of such a system, we are therefore interested in a modular approach: an overall design structure that can be developed and improved with new technological solutions, better sensors, etc.: in other words, a sustainable design for future scenarios. In this paper we present an attempt toward such a design, as well as the first steps of implementing a VE-based teleoperation system for a forestry forwarder.

2. System Design: Scenarios and Requirements

Scenarios can be used in many ways during a design process. By using scenarios of VE usage in a forestry machine setting we can concretize our visions about the future and study unrealized ideas from a user perspective. At this stage, we are working with conceptual scenarios. This means we want scenarios that are concrete enough to capture essential parts of the interaction and contain many of the fundamental requirements for the system. At the same time they should be abstract enough to not limit our choices, e.g., of interaction methods or implementation strategies.

Due to the long-term time perspective we do not focus on the specific technological aspects and solutions. The rate and direction of technological research and product development cannot be predicted and the availability of new products may change some of the more low-level aspects of the design. Therefore, we strive toward a system design that is independent from specific technology choices until a later stage in the process.

We have selected three different scenarios as a basis for designing a VE as well as for evaluating its usefulness in different stages of technological advancement and autonomy. The three scenarios are our attempts to answer three different categories of questions:
1. What can be achieved without environment recognition? How can the operator be assisted?
2. What functionalities and opportunities does the environment recognition bring? How does it change the use of the system and the role of the operator?
3. What can the teleoperation aspect contribute with? What are the potential benefits and difficulties? How does remoteness affect the requirements for the system?
The scenarios will be used to extract functional requirements for the system. Similarities and differences between the interaction methods in the different scenarios will shape the structure of the HRI system.

2.1 Scenario 1: Operator Assistance

In this scenario the operator is placed in the cabin of the forestry machine, equipped with semi-autonomous capabilities and a VE user interface designed to assist the operator. The operator could also work from nearby the machine using on-site teleoperation, i.e., with direct visual overview of the scene. Sensors continuously acquire information regarding the configuration of the crane, allowing an up-to-date representation of the crane and the state of each of its links (Figure 2). This, as well as other relevant information, e.g., crane trajectory representations, is displayed in a virtual environment and presented to the user in an unobtrusive way.

The manual crane operation is still a large part of the work, however the operator has access to a selection of automated trajectories that can be automatically executed on command. The VE system is linked to an interface to specify a target point for the end effector, appropriately selected (by the user) to be in the vicinity of where a task should be performed, e.g., near a log that should be grasped.

No real-time information about the environment is presented, i.e., objects in the workspace of the vehicle are not apparent in the VE. Even if the surroundings are not visualized in the VE, the real-time tracking of the crane position and the known spatial location of the trajectories allow the user to get visual feedback on the progress of the motion. This also lets the operator plan the next actions to be taken (instead of being occupied by controlling the motion) and thereby reduces the stress and mental load.

Additional useful features that the VE can provide is a global map, which shows the borders of the current cutting area, roads, terrain information, etc. This can help the drivers to organize and plan their tasks or even add information to other drivers, e.g., recommended driving paths.
2.2 Scenario 2: Environment Awareness

For this second scenario, we add the possibility of showing a detailed representation of the local environment in the VE. This means that objects near the vehicle are sensed, recognized and classified, and dynamically added to the VE.

A similar high-level control input method is used (as in Scenario 1), with the major difference being that here the selectable trajectories can be adapted to the locations of the trees or logs that are to be manipulated. This means that instead of a target position, the user uses an intuitive and efficient user interface to select a target location or a target object, an object of interest (OOI), as visualized in the VE. The system uses the OOI, along with the other objects in the VE, to plan an appropriate trajectory that avoids obstacles.

The view of the operator in a real machine is restricted. Consequently, the virtual representation of the physical environment around the machine is used by the operator to get an overview of the surroundings, and it can in some basic ways assist in the more high-level planning (in which order to pick up logs, etc.) and visualize these plans.

2.3 Scenario 3: VE Teleoperation

In this most futuristic scenario we consider off-site teleoperation of the forestry machine, i.e., no direct line of sight between the operator and the machine. In order to reduce the workload for the operator, automated control is conducted when considered feasible, complemented with some elements of direct control. Due to robust and reliable sensors, as well as a machine that is designed for automation, the role of the operator is more supervisory. High-level commands are given to the system through an intuitive interface: motion planning and execution with obstacle avoidance is performed autonomously while a virtual environment gives feedback to the operator. An exocentric view helps the operator to plan the high-level tasks, while an egocentric view can be used for instances of direct control. Multimodal feedback presents important information to avoid overload of the visual-sensory channel.

2.4 Requirement Analysis

The next step in the design process of the interaction and control system is to move from the more general and conceptual ideas that the scenarios represent to more functional details in order to extract the requirements.

For the first scenario, the focus is on semi-automation, i.e., on implementing efficient automated crane trajectories and providing an intuitive user interface that helps the operator to select and execute these motions.

By design, the largest difference between the first two scenarios is the environment reconstruction system. For the second scenario, not all objects must be classified, but the system should detect OOIs and classify other objects into basic categories that cover their relevant properties. For instance, it is desirable that the system is able to distinguish between real obstacles (e.g., rocks and trees) and “soft” material that can be disregarded (e.g., grass and bushes). Since the operator still has access to a real-world view of the workspace, there is no need for a high level of detail of the virtual representation; instead, simpler abstractions that are easy to recognize can be used. The reconstruction also opens up to a more intuitive interaction, as well as more advanced automation (involving decision making) that in other ways already requires environmental knowledge. Although not necessary, such functionality will be included in the requirements as desired features.

In the third scenario, the operator must rely solely on the information given by the sensors and interpreted by the system. This imposes high demands on the reliability and robustness of the environment recognition system—more so than in the previous scenario, where the operator was presented with redundant input about the crane and environment (VE as well as direct vision by the
Table 1: Requirements with priority levels (H=High, L=Low) for each scenario (S1–S3).

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hardware</strong></td>
<td></td>
</tr>
<tr>
<td>Sensors for all configuration variables</td>
<td>H</td>
</tr>
<tr>
<td>Sensors for environment recognition and classification</td>
<td>H</td>
</tr>
<tr>
<td>Input devices no less ergonomic than currently used methods</td>
<td>H</td>
</tr>
<tr>
<td>Output devices that can present a VE in an unobtrusive way</td>
<td>L</td>
</tr>
<tr>
<td><strong>Environment reconstruction</strong></td>
<td></td>
</tr>
<tr>
<td>Detection of logs, trees, and other objects of interest</td>
<td>H</td>
</tr>
<tr>
<td>Detection of obstacles in the workspace of the manipulator</td>
<td>L</td>
</tr>
<tr>
<td>A reliable and robust environment recognition system</td>
<td>L</td>
</tr>
<tr>
<td>Reaction system for detection of humans in the work area</td>
<td>H</td>
</tr>
<tr>
<td><strong>Motion planning and control</strong></td>
<td></td>
</tr>
<tr>
<td>A procedure to find efficient trajectories connecting target points</td>
<td>H</td>
</tr>
<tr>
<td>A control method to efficiently and reliably execute a selected trajectory</td>
<td>H</td>
</tr>
<tr>
<td>On-line motion planning</td>
<td>L</td>
</tr>
<tr>
<td>Obstacle avoidance</td>
<td>L</td>
</tr>
<tr>
<td>Manipulation of predefined target points or trajectories</td>
<td>L</td>
</tr>
<tr>
<td><strong>User interface</strong></td>
<td></td>
</tr>
<tr>
<td>Task-oriented user interface for semiautonomous elements</td>
<td>L</td>
</tr>
<tr>
<td>Visualization of current crane configuration and status</td>
<td>H</td>
</tr>
<tr>
<td>Present the operator with a good overview of the local surroundings</td>
<td>H</td>
</tr>
<tr>
<td>Present the operator with an overview of the global working area</td>
<td>L</td>
</tr>
<tr>
<td>Assistance in high-level planning</td>
<td>L</td>
</tr>
<tr>
<td><strong>Information presentation</strong></td>
<td></td>
</tr>
<tr>
<td>Detailed, occlusion-free view for free-hand manipulation</td>
<td>L</td>
</tr>
<tr>
<td>Easily accessible exocentric view for high-level planning</td>
<td>L</td>
</tr>
</tbody>
</table>

operator). More input may, however, give the operator more to focus on and hence lead to a higher mental workload. For safety reasons, the operator may need to also have access to raw data such as video streams in order to ensure the integrity of the information in the VE scene. This should also be complemented with some human and animal detection system that is robust and reliable.

Our compilation of requirements derived from the scenarios is shown in Table 1. For each scenario, the relevant requirements are labeled as necessary (H = high priority) or desired (L = low priority).

The three basic scenarios illustrate some of the flexibility of a VE system by describing some different ways that ordinary work cycles can be improved by including a VE system into the human-machine interaction. Figure 3 illustrates how a VE can be integrated as part of the information flow in a forestry machine system at different stages of the automation process.
Figure 3. Information flow in the different scenarios. Bold text and lines indicate details that are scenario-specific (i.e., not common to all scenarios).
3. System Implementation and Evaluation

Using the three scenarios and the analysis of the resulting requirements as a starting-point for the design process, a prototype system has been constructed, intended as an experimental and evaluation platform for human-machine interaction during forestry machine operation. The system includes VE software for visualization and user interaction, and it has been used in conjunction with a forwarder crane equipped with prototyping hardware.

3.1 Experimental Setup

The following experimental equipment has been installed at a laboratory at Umeå University:

- Cranab 370RCR hydraulic forwarder crane,
- Komatsu operating chair with joysticks and controls,
- dSPACE MicroAutoBox real-time prototyping hardware, with an on-board PowerPC CPU,
- RapidPro power amplification unit and signal conditioners,
- Three joint-position encoders,
- Retracting wire sensor to measure telescope extension,
- Magnetic sensor for grapple rotation (P. X. La Hera, Mettin, Westerberg, & Shiriaev, 2009),
- Hydrotechnik transducer HD pressure sensors (used for modeling and control),
- Point Grey Bumblebee2 stereo camera,
- Structured light camera (640x480 resolution for stereo and RGB image).

The 370RCR crane is somewhat smaller than most cranes used on real forwarders but was chosen due to space restrictions in the laboratory. Its operational principle and dynamics are, however, very similar. The main experimental setup is shown in Figure 4. However, some of the experiments have been performed on a commercial machine, a Komatsu 830.6 forwarder, using the same type of electronics and sensors.
3.2 System Architecture

Figure 5 shows the architecture of the VE-assisted crane control system. The system consists of a number of loosely connected subsystems with separate responsibilities. The modularity of the system allows for replacement of user interface technologies, algorithms for analyzing sensor data, and crane control methods, with only small adjustments to the overall system structure.

At the center of the system is the CraneVE component, which is responsible for visualization and data communication between other modules. All inputs and outputs to the CraneVE module are designed to be generic by managing high-level objects and tasks. Middleware-like software components take care of transforming sensor data to recognizable and classified objects.

Communication between modules is performed using UDP/IP, which allows remote control over either a local area network (LAN) or over the Internet.

The modularity also allows for distributed visualization and control. Sensor modules can send data to different virtual environments, and the same VE can be used to send commands to different machines. This can be useful for supervision or high-level planning.

Each module will be described individually in the following sections.

3.3 Software

All the software for motion planning and control was developed using Matlab and Simulink. The communication between the host PC and the dSPACE real-time processor gets mediated through a host computer, running client-server architecture software using the dSPACE CLIB interface. Matlab and Simulink is also used for the user input functionality.

CraneVE was implemented in C++, using the open source graphics toolkit OpenSceneGraph to structure and visualize the graphical data.

The algorithm for log recognition was implemented in C++, using the computer vision library OpenCV for some of the image manipulation tasks and OpenGL for visualization of the results.

3.4 Motion Planning and Crane Control

The workspace for a large (10m) forestry machine crane is extensive. However, most of the performed crane motions have similar shape and the active workspace is usually concentrated within a smaller region. This means that only a limited range of motions must be implemented in order
to achieve useful automation functionality. Typical target points for the end effector of a forwarder, given by analysis of recorded data from manual operation, are shown in Figure 6. Relatively few target end points would be enough to cover a large part of this active region. When environment recognition is unavailable, these points could be used as destinations for which preplanned motions are accessible for the driver through some kind of interface. A large number of feasible and optimized motions can be constructed in advance (off-line planning) and stored in a database. Where the desired motion could not be anticipated, e.g., when obstacles are present, an automated motion can still be performed through real-time planning.

3.4.1 Off-line Motion Planning Motions that are preplanned and selectable by the user should preferably be fast and energy efficient. Using the motion planning method of virtual holonomic constraints (Mettin, La Hera, et al., 2009), optimal trajectories along some desired path can be calculated for a suitable performance index.

First, a reasonable path from the start point to the end point is decided, either by adaptation of relevant recorded motions performed by operators or, preferably, using an optimization procedure. With a given path, the speed and relative use of the different crane links along this path can be shaped to construct a trajectory with optimal performance within the limitations of the machine. Velocity constraints of the individual joints are particularly restrictive in hydraulic manipulators. The motions are then tracked using feedback linearization. Mettin, La Hera, Morales, Shiriaev, Freidovich, & Westerberg (2009) have described this process in detail.

In Mettin, Westerberg, Shiriaev, & La Hera (2009), trajectories planned with this approach are compared with recorded motions by professional human operators. In the study, the focus is on optimizing trajectories for execution time. Results show that performance can be significantly improved by path-constrained re-planning of human-operated motions, i.e., following the same Cartesian path, but with a redistribution of work between the different joints of the crane (allowed by the redundant design of the manipulator). Additionally, trajectory re-planning for time-efficient geometrical paths along with efficient velocity profiles along the paths can further improve the time efficiency of the crane motions.

Figure 6. Some examples of trajectory end points for the end effector of a forwarder crane, as recorded from actual crane work.
3.4.2 Real-time Motion Planning and Control Since some motions cannot be anticipated in advance, like the ones created by a collision avoidance algorithm, they need to be efficiently constructed in real time. Our implementation of the virtual holonomic constraint method does not allow motion planning in real time. The problem can be solved by combining shorter preplanned motion segments; this is, however, complicated (since it requires matching initial and final velocities where the segments are joined) and the resulting combined motion does not conserve the optimality property of the individual segments. In our case, we instead use a method based on pseudo-inverse Jacobian control for real-time planning that is standard in robotic applications. This method is described in further detail in Westerberg, Manchester, Mettin, La Hera, & Shiriaev (2008).

3.4.3 Human-Machine Collaboration An important part of a semi-automated system is efficient collaboration between the human operator and the computer control system. The details of such collaboration could depend on the state of the operator, e.g., their current workload or focus of attention, as well as on aspects of the environment, such as obstacles or unexpected events that need immediate concern. In the most basic form of human-machine collaborative work, the collaboration takes place on the control level by defining the effective control input to the system as some function of the individual control efforts. Depending on how this function is selected, the division of labor is usually categorized into either shared or traded control (Sheridan & Verplank, 1978; Sheridan, 1992). In shared control, the human and computer actors work simultaneously on the same task. For instance, the human operator can apply real-time adjustments to an autonomously executed motion in order to avoid obstacles that are unknown to the computer control system. Traded control, on the other hand, divides the tasks in the temporal domain, i.e., the human and computer either takes turns working on the same task or performs different tasks that are temporally interdependent. In the scenario of forestry crane control, crane motions from the load bunk toward the logs could be automized, while the human operator controls the grasping and loading manually.

In Hansson & Servin (2010), shared and traded control was investigated using the hardware and general software framework of the crane laboratory setup. From the experimental results it could be shown that the performance for inexperienced users is largely increased when provided with either of the semi-automation strategies. Experienced operators performed slightly worse with shared control, but considering that they have several years of training in manual control, it is expected that extended training as well as fine tuning of the motion planner and individual link controllers could level this difference. Integrating traded control with the off-line motion planning method would likely increase the productivity beyond that of current manual control. Further, both categories of users considered the semi-automated tasks to reduce the cognitive workload and to allow the user to change focus from low-level crane control to other functions, such as more high-level task planning.

3.4.4 Obstacle Avoidance One of the advantages of using a VE system with external information is that it can assist by planning a collision-free path around obstacles. This could be implemented as an on-line modification of a path preplanned with the virtual holonomic constraint method as described above. Another way is to construct and follow a new collision-free path using real-time motion planning. Our implementation uses the latter method and features a simple obstacle avoidance algorithm (as described in Figure 7). Objects in the VE are automatically considered obstacles and taken into account when the crane trajectory is planned. An example of a path constructed by the algorithm can be found in Figure 8.

3.5 User Interface and Operator Assistance
The user interface is an important part of the VE system, and a number of options are available for user input technology. Apart from using joystick motions reconfigured to fit the changed functionality, it is in theory possible to use a range of technologies, e.g., voice recognition, eye tracking,
Figure 7. The collision avoidance algorithm. (a) Assume that $p$ is the end point of the current path. Let $q$ be a new target point defined by the operator. (b) If $q$ collides with an object, find a safe point $q'$ above the bounding sphere of the object. (c) If the path $p \rightarrow q'$ collides with an object, find a safe point $c'$ above the center of the object. Let $m$ and $n$ be the middle point of each sub-path $p \rightarrow c'$ and $c' \rightarrow q'$ respectively. (d) Find safe points $m'$ and $n'$ above $m$ and $n$ such that they are outside of the bounding sphere of the colliding object. The final path will be $p \rightarrow m' \rightarrow c' \rightarrow n' \rightarrow q'$.

gesture recognition, and 6DOF input devices. This being said, using a more intuitive 3D user interface for the non-automated manipulation tasks or for adjusting preplanned trajectories can still be very useful and may be both easier to learn and more efficient than today’s individual link control.

In our implementation, we have chosen to use a regular joystick as input, but instead of simultaneous use of both hands, we allow the crane motion planning to be performed with only one joystick. Thereby only one hand at the time is needed, which even more increases the relaxing time that is already introduced by the autonomous motions.

The user can either select from a set of predefined target points, use the joystick to select a custom 3D target point, or select a predefined point and use the joystick to make corrections of the path by moving the end point while the crane is moving. The predefined points can be constructed from knowledge about common motion endpoints, as in Figure 6, or according to user preferences.

An example of selectable targets with associated trajectories is shown in Figure 9. The range of target point options is adapted to the current position of the crane, e.g., if the crane is extended, its next target is likely to move grasped logs to the load bunk (Figure 9c). When using the joystick to
select target point, the position will be initialized at the target point for the previous trajectory of the same category, taking advantage of the fact that consecutive trajectories often are similar and end at similar positions.

In scenario 2 and 3 where the environment is recognized and visualized in the VE, target points can be adapted to the logs or other OOI that are located in the workspace of the crane. In this case neighboring OOI are clustered into groups, in order to avoid cluttering of the scene (Figure 9d).

3.6 Environment Reconstruction

In order to be able to fulfill the requirements from the last two scenarios, we need the possibility to show a detailed representation of the local environment in the VE. This includes recognizing and classifying objects and dynamically updating the VE.

Recognition of unknown and unstructured outdoor environments in general is a difficult task. The complexity of the object classification task generally increases with the number of object classes that need to be correctly classified. In our case of automation of crane motions however, only a few classes of objects are of interest for the system, e.g., logs that should be picked up or objects that affect the motion. Some of these are known, such as different parts of the vehicle, and others are unknown, such as trees and rocks. In the CraneVE system a visualization and communication API has been implemented for the categories of tree, log, human, and obstacle (represented as a convex hull). Each object can be associated with properties like position, orientation, size, and color.

Even though detecting obstacles and classifying terrain for vehicle navigation purposes is useful, and even necessary for Scenario 3, trees and logs are most important for Scenario 2. An experimental method for classification and detection of logs has been developed (Park et al., 2011) where the resulting virtual objects are shown in the VE and can be used to define target points. A structured light camera is used as data acquisition hardware (Figure 10a). The camera uses a light source to project a pattern of light that is recorded by an RGB camera (Figure 10b). Based on the transformation of the reflected light, depth values for each pixel can be calculated. This 3D data along with information from the acquired color image is used to find cylindrical-shaped objects with tree-like...
color and texture. An example of detected trees is shown in Figure 10c with the corresponding result in the VE shown in Figure 10d. This specific technology was selected due to a good 3D point cloud quality, simplifying the design stage of the detection algorithm. This particular hardware is, however, not built for practical real-world use due to sensitivity to lighting conditions and limited range. The current implementation of the algorithm also does not work in real-time.

Evaluation of the environment reconstruction system was performed with respect to the success rate of reconstruction in different configurations. A number of logs of various diameters were arranged in different orientations $[0^\circ, 30^\circ, 45^\circ, 60^\circ, 90^\circ]$ and distance intervals $[2m - 5m]$. In total, 237 test images were acquired in illumination $[200 - 500\,lx]$ using our structured light camera.

The success rate of the log reconstruction is presented in Table 2. The point cloud error is proportional to the distance between the camera and the object, as reflected by the results. With these particular equipment and algorithms, reliable log recognition is not achievable for the longest distances required in a real-life scenario. However, the results are reasonable at shorter distances; for correctly recognized logs, both the position and the size are estimated accurately enough, which encourages further research and development of the method.
3.7 Teleoperation

By separating the system of Figure 5 into two subsystems residing on different physical locations, teleoperation is achieved. A teleoperation system based on CraneVE is described in Figure 11. The architectural difference to the system presented in Figure 5 is merely conceptual, since all the subsystem communication already exists and can be executed over the Internet.

The separation does add further requirements on the robustness and reliability of the environment sensing system since the operator would have no possibility to compare the VE with the real environment. It also introduces time delay to the control loop. The length of the delay depends on the communication distance and possibly on the available bandwidth. Note, however, that in our realization local controllers for autonomous preplanned motions are used to avoid performance deterioration due to delays.

Teleoperation functionality is fully implemented and has been successfully demonstrated for long-distance remote control. We have also tested the real-time control and motion planning implementation, and overall, we were able to make all links converge to the desired positions with almost zero steady-state errors, sufficient speeds, and without oscillations. An extended description of the teleoperation specific control algorithms and experimental results can be found in Westerberg, Manchester, Mettin, La Hera, & Shriaev (2008).
Table 2: Success ratio of log reconstruction with respect to distance.

<table>
<thead>
<tr>
<th>Distance</th>
<th>Success rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2m</td>
<td>84%</td>
</tr>
<tr>
<td>3m</td>
<td>65%</td>
</tr>
<tr>
<td>4m</td>
<td>58%</td>
</tr>
<tr>
<td>5m</td>
<td>33%</td>
</tr>
</tbody>
</table>

4. Partial Concept Verification: Experimental Results

The main focus of the three proposed scenarios is the virtual environment (VE) and the automated crane trajectory planning. As successful results in path planning and control have been presented in earlier studies (P. La Hera, Mettin, Manchester, & Shiriaev, 2008; Mettin, Westerberg, Shiriaev, & La Hera, 2009; Morales et al., 2011), we will now consider the VE, which majorly influences both the system requirements and the modular design and implementation. It is therefore also the most important element to consider when evaluating the feasibility and effectiveness of the scenarios and the resulting requirements-based design. We have accordingly designed an experiment to independently test the effects of VE feedback on different performance aspects. In this experiment we compare two different VE visualizations with traditional direct view as well as with video camera feedback (the competitive method for visual feedback in teleoperation).

4.1 Materials and Method

For the comparative user study we engaged 17 test subjects. All participants were male, ages ranging from 22 to 31 (M = 25, SD = 2.74). None of them had previous experience with forestry machine control.

The participants were asked to navigate the crane between logs on the ground in a given order while avoiding obstacles. The grapple was to be positioned above the center of each log and rotated to align with the log as if picking it up. In order to decrease learning time, a simplified control interface was used. Instead of regular control, where two joysticks are used to control the individual links of the crane, the position of the crane was controlled using a single joystick. Specifically, the joystick position was mapped to the rate of change of variables in a polar coordinate system \((\theta, r)\), where \(\theta = q_1\) represents the slewing and \(r\) is the extension, i.e., the distance from the crane base.
The other joystick was used to control the rotation of the grapple. The grapple itself was not used during the experiment.

The task was repeated for different modes of visual feedback. The following methods were considered:

1. **Direct View (DV)**: On-site operation with direct visual view of the work area.
2. **Video Camera (VC)**: Simulated teleoperation with video camera feedback.
3. **Virtual Environment, regular view (VE1)**: Simulated teleoperation with virtual environment feedback, replicating the view of the video camera.
4. **Virtual environment, top-down view (VE2)**: Simulated teleoperation with virtual environment feedback, task-adapted top-down view point.

For the video camera feedback method, a Sony HDR-CX115E camcorder with 1080p HDMI output was used to capture the workspace of the crane. The camera view is designed to replicate the view point of the direct-view condition, though it naturally presents a more limited field of view. The video signal was displayed on a 24” LCD monitor. The same monitor was used to display the virtual environment views. The regular view presented in VE1 displays the virtual scene from a similar view point as the video camera. In VE2, the camera view is adapted to the task by providing a dynamic top-down view with a virtual camera that follows the motions of the crane.

After a training session to get familiar with the controls and the dynamic behavior of the crane, the participants performed the task for each of the conditions in a random order. The dependent variables in the experiment were task completion time and accuracy in target position estimation. In addition, we recorded the number of collisions occurring between the crane and the obstacles.
Each participant was also asked to complete a questionnaire regarding the experienced difficulty level of individual elements of the task, as well as the general impressions for each visual feedback method. They were also asked to order the methods in terms of personal preference. Each test took approximately 45 minutes.

4.2 Results

4.2.1 Task completion time The effect of the visual feedback methods on task completion times was analyzed with a repeated measurements ANOVA test. Results showed a significant effect, $F(3, 51) = 13.51, p < .05$. The data distribution for the task completion times is shown in Figure 13. Significant differences revealed by post-hoc analysis (dependent-sample t tests with Bonferroni correction) are indicated by asterisks. The task completion time is lower for the DV interface as compared to VC, $t(17) = -3.52, p < .003, r = 0.65$, and VE1, $t(17) = -4.88, p < .001, r = 0.76$. For VE2, the completion time is significantly lower than for VE1, $t(17) = 4.69, p < .001, r = 0.75$.

4.2.2 Accuracy The accuracy of the crane positioning was measured for each target. Each time the participant pressed the button to indicate a complete motion, the crane position was registered and compared to the true position of the target. The average error between the true and measured positions was used as dependent variable. A repeated-measures ANOVA test showed a significant effect on mean error levels, $F(3, 51) = 14.99, p < .05$. Figure 14 shows the distribution of the crane position error for each test condition. Post-hoc pairwise comparisons using dependent samples t tests with Bonferroni correction showed significantly higher accuracy (lower distance error) for VE2 as compared to both DV, $t(17) = 4.65, p < .05, r = 0.75$; VC, $t(17) = 6.39, p < .05, r = 0.84$; and VE1, $t(17) = 4.33, p < .05, r = 0.72$.

Another measure of the effects of visual feedback method on crane control accuracy is the number of collisions between crane and obstacles. The total number of collisions for all participants is shown in Figure 15. Most collisions occurred using the video camera feedback.

4.2.3 Subjective evaluation In the post-experiment questionnaire, the participants were asked to estimate the workload for each of the visual feedback methods (ranging from 1=low to 7=high). The results are shown in Figure 16.

The participants were also asked to evaluate the difficulty of the following specific elements of the task:
1. Estimating the exact position of the crane in relation to the targets and obstacles,
2. Navigating to the target in a smooth motion without interruptions,
3. Rotating the grapple to the correct orientation,
4. Simultaneously controlling crane position and grapple rotation.

The answers were given on a scale between 1 (easy) and 7 (hard). The results are shown in Figure 17.

Finally, the subjects were asked to rank the user interface methods by preference. For this purpose, they were asked to order each of the feedback methods using the numbers 1 (most preferable) to 4 (least preferable). The results are given in Figure 18.

Friedman’s ANOVA revealed a significant effect of the feedback method on the preference ratings, $\chi^2(3) = 25.2, p < .05$. Post hoc testing using Wilcoxon comparisons showed that direct view (DV) had significantly higher user preference than video camera (VC), $z = -3.11, p < .05, r = 0.52$, and regular VE view (VE1), $z = -2.94, p < .05, r = 0.49$. A similar significant difference was shown for the top-down view (VE2) as compared to VC, $z = -2.86, p < .05, r = 0.48$, and VE1, $z = 3.15, p < .05, r = 0.52$. 

103
Figure 13. Task completion times using direct view (DV), video camera (VC), virtual environment with regular view (VE1), and top-down view (VE2).

Figure 14. Average distance error between the crane tip and the intended target positions.

Figure 15. Total number of collisions for each condition.
Figure 16. Perceived workload for performing the task using direct view (DV), video camera (VC), virtual environment with regular view (VE1), and top-down view (VE2).

Figure 17. Perceived difficulty level of some individual elements of the task.
4.3 Discussion

The virtual view presented by the VE1 condition was designed to closely match the video camera view. Even so, a few relevant differences can still be found that should be taken into consideration when analyzing the results.

Firstly, the orientation of the grapple is obtained using a magnetic encoder mounted on the rotator. The low resolution (10 degrees) of the sensor leads to a significant quantization error. Consequently, VE users will receive infrequent quantized updates of the grapple orientation. In general, the time between the updates depends on the rotational speed of the grapple, which makes the effect most noticeable during finer rotator adjustments near the target orientation. Since novice users tend to rely heavily on visual feedback for manipulation tasks (as opposed to motor memory for more experienced users), the outdated visual information can, in these situations, easily lead to overshoots where the user repeatedly overestimates the required control effort. This will have a negative impact on the task completion times.

Secondly, video camera feedback (as well as direct visual feedback) also gives access to more detailed information about the dynamics of the crane motion. In particular, small oscillations in the unmeasured and passive degrees of freedom of the grapple may introduce disturbances in the position control system and introduce system-wide oscillations. These disturbances may be harder to counter for the operator using VE feedback, since their origin is hidden. This could have a negative impact on both speed and accuracy for the VE methods.

Despite these differences in the presentation of information, however, no significant differences could be found between the VC and VE1 conditions for the time and accuracy measurements. Of the three remote view conditions, the participants performed best using the task-adapted virtual environment (VE2). The effect was largest for accuracy, both when measured with distance error and number of collisions. This result supports the proposed hypothesis that the adaptability of the VE makes it a superior method for visual feedback in forestry crane teleoperation.

When comparing the results from the direct visual feedback condition with the teleoperation simulating counterparts, it is important to note that no stereoscopic display was used for the VC

![Figure 18](image)

*Figure 18.* Subjective ranking of the different user interface methods, direct view (DV), video camera (VC), virtual environment with regular view (VE1), and top-down view (VE2).
or VE scenarios. Only 2D depth cues, such as perspective and shadows, were provided to the participants. The limited depth perception from the lack of stereopsis in the current implementation presents a disadvantage to the simulated teleoperation methods. The field of view (FOV) is also smaller for a screen than for direct view. A $24''$ screen at $75cm$ distance translates to a diagonal FOV of approximately 45 degrees. Consequently, only a portion of the crane linkage is shown on the screen for the VC and VE1 conditions, which may further limit depth perception as well as have negative influence on the user's mental model of the crane behavior. In addition, the differences mentioned between VC and VE apply also when comparing DV and VE.

Despite these limitations, performance with virtual environment feedback was high even compared to the direct view condition. Though the mean performance for VE2 was slightly worse in task completion time, the difference was not significant. In accuracy, VE2 instead yielded significantly better results than DV with the mean distance error of about 50 percent higher for DV. Both DV and VE2 were given favorable ratings on perceived difficulty levels, and both were generally selected as first or second choice in terms of personal preference with no significant difference between the two conditions.

It is not trivial to transfer these results to a current real-world scenario and even less so for a future scenario where the operators’ tasks may be heavily modified. For this specific task however, it is shown that the VE can be adapted to better present the characteristics that are important for an accurate positioning of the crane. For manipulation tasks, there is generally a tradeoff between task completion time and accuracy. Additional practice would let the user become more aware of how to balance the relative importance of time and accuracy in the crane motions. It is possible that for motions performed with an accuracy level matching the direct-view condition, the task completion time could even be lower using the VE. Further, a freely adaptable camera view is only one of the possible task dependent operator assisting features that a VE can provide. Other visualization or interaction methods, e.g., predictive display, haptic feedback, or visual object categorization for high-level planning, etc., could be used to improve results also for other performance measures. It is also expected that the subjective ranking of a VE interface would improve when more of these features are implemented and the advantages of the VE are more apparent. One example of such advantage could be the use of a fully customizable viewpoint, which gives the operator an overview that is not possible using a direct-view approach without leaving the cabin.

5. Concluding Remarks

We have presented initial steps toward the introduction of teleoperation in industrial forestry machines.

Relevant conceptual scenarios were constructed to concretize visions about future interaction from a user perspective. An analysis of the scenarios resulted in a list of requirements, each of which was assigned a priority for each of the different scenarios. In the design process, this allowed us to find a common basis to build the system upon and to create a modular design that lets us add additional functionality for each scenario.

Even though scenarios are useful as tools to design interaction systems oriented toward specific goals, the work with novel interaction methods can also let us discover new uses for the system components, e.g., how a VE system could be useful outside of the daily forwarder operation. Many of the suggestions presented in the scenarios can be particularly useful for operators in training, where evaluation of crane control technique and general performance can be presented along with suggestions for improvements in motion planning and execution. Forestry machines can be monitored in real-time and operators’ crane trajectories can be logged and analyzed visually as well as statistically, e.g., for understanding how to improve the crane motions. Furthermore, the VE system can be used for large-scale planning in order to optimize production or minimize negative environmental
effects. Letting novice operators teleoperate a machine in a safe training area can be a good complement to simulator training, since the introduction of hardware into the simulation loop will affect the operation experience. Desired motions or effects from performed motions can be visualized and operators can learn the optimal motions by visualization and guided assistance from force feedback joysticks.

As a product of the design process, we have implemented a prototype virtual environment teleoperation system. This system is not intended to be viewed as something that can be fully used today; for much of the functionality presented in the last two scenarios, the technology and methods needed for sensing the environment would need to display a degree of reliability that is not yet available. We have, however, performed successful tests of conceptual ideas aimed to be integrated with the results from our ongoing work in motion planning and control methods for semi-automation. Through user tests we have demonstrated some of the advantages of VE feedback for teleoperation regarding view adaptation for improved accuracy as well as a reduction of the perceived workload. The results are not complete but there is enough data within our study to encourage further development.

A number of challenges still exist that are relevant to consider in future work. For environment reconstruction, there is a need for robust sensor technology that can provide high quality information as well as withstand the rough conditions of the outdoor environment. Appropriate user interface methods need to be investigated and evaluated. Additionally, time delay in communications has not yet been considered. These are a few of the challenges that must be faced before we can reach the end goal of semi-autonomous and teleoperated forestry machines. However, implementation of the first steps in commercial machines need not be far away and the results of this design process will be further developed and used as an experimental platform for further research regarding human-machine interaction in forestry and field robotics.

Whenever the individual parts have been integrated into a semi-autonomous system that allows smooth cooperative work over the entire work flow, additional aspects of the user interface methods should be investigated and evaluated. In particular, effects on the physical and cognitive workload should be considered, e.g., how semi-automation affects the performance during a longer work shift as compared to fully manual control. Additionally, it is of interest to see how the specific implementation of the user interface and the collaborative actions affects the overall performance of the system.

References


Hannaford, B. (2000). Feeling is believing: History of telerobotics technology. In The robot in the garden:
telerobotics and telepistemology in the age of the internet (pp. 246–274). Cambridge, MA, USA: MIT Press.


Authors’ names and contact information: Simon Westerberg, Department of Applied Physics and Electronics / Industrial Doctoral School, Umeå University, Umeå, Sweden, simon.westerberg@umu.se; Anton Shiriaev, Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway, anton.shiriaev@itk.ntnu.no.