# AntHand: Interaction Techniques for Precise Telerobotic Control Using Scaled Objects in Virtual Environments

Dries Cardinaels
UHasselt - Flanders Make
Expertise Centre for Digital Media
3590 Diepenbeek, Belgium
dries.cardinaels@uhasselt.be

Bram van Deurzen
UHasselt - Flanders Make
Expertise Centre for Digital Media
3590 Diepenbeek, Belgium
bram.vandeurzen@uhasselt.be

Raf Ramakers
UHasselt - Flanders Make
Expertise Centre for Digital Media
3590 Diepenbeek, Belgium
raf.ramakers@uhasselt.be

Kris Luyten
UHasselt - Flanders Make
Expertise Centre for Digital Media
3590 Diepenbeek, Belgium
kris.luyten@uhasselt.be
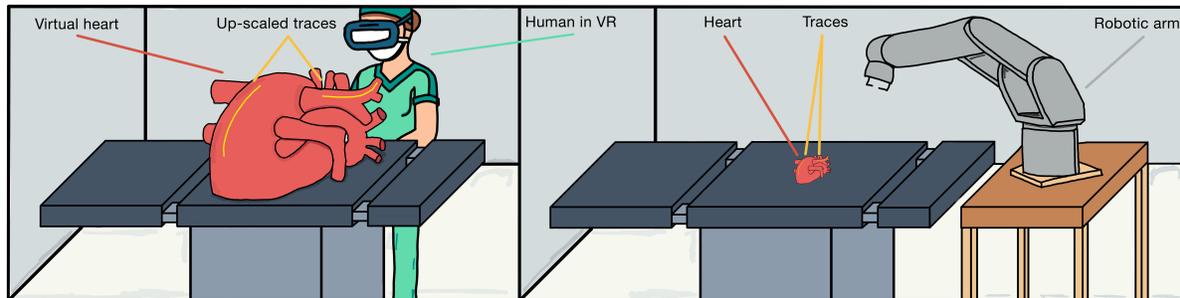
Figure 1: In the left part, a virtual environment contains a scaled heart being manipulated by a surgeon: yellow traces indicate the manipulation done in the virtual environment. On the right side, the physical environment is depicted with a life-size heart, accompanied by yellow traces indicating the manipulations carried out by the robotic arm.

ABSTRACT

This paper introduces AntHand, a set of interaction techniques for enhancing precision and adaptability in telerobotics through the use of scaled objects in virtual environments. AntHand operates in three phases: up-scaling interaction, for detailed control through a magnified virtual model; constraining interaction, which locks movement dimensions for accuracy; and post-editing, allowing manipulation trace optimization and noise reduction. Leveraging a use-case related to surgery, the application of AntHand is showcased in a scenario demanding high accuracy and precise manipulation. AntHand demonstrates how collaboration between humans and robots can improve precise control of robot actions in telerobotic operations, while maintaining the familiar use of traditional tools, rather than relying on specialized controllers.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; **Interaction design theory, concepts and paradigms**.

KEYWORDS

Human-Robot Interaction; Virtual Reality; Telerobotics; Precise Interaction

1 INTRODUCTION

Robots are used in many applications requiring high precision, speed, and repeatability [4, 6]. However, they encounter limitations in adaptability, interaction with humans, and programmability mainly due to their focus on working in static environments within safety enclosures [5]. In response to these limitations, Collaborative robots (cobots) have emerged, designed explicitly for direct human interaction [9], increasing adaptability and flexibility by enabling

humans and robots to work together. However, as part of their inherent safety, cobots are limited in load and speed.

Telerobotic collaboration enables combining the flexibility and adaptability of humans with the precision, speed, and load of high-precision robots. Precision and ease of operations can be further improved through interaction paradigms such as shared control [3]. Through shared control, precision input from operators is combined with sensory feedback from the robot. While the sensor system can intervene in inaccurate situations, the robot's precision remains intricately connected to the operator's movements in a shared control context. Additionally, the precision of the operator in these systems is also impacted by the accuracy of the visual information presented [10].

Surgical robots offer an interesting combination between a cobot and a teleoperated robot, as they are controlled by a surgeon but used in the same environment as the medical staff during the operation. A surgical robot combines a robot's high precision with a surgeon's knowledge and expertise to execute complex operations. Surgical robots, however, require special controls that are different from traditional medical tools and thus require additional training.

In this paper, we present AntHand; a set of interaction techniques for teleoperating robots with high precision, exemplified through a surgical procedure on a heart. Unlike traditional techniques for teleoperating robots, our interaction technique does not require specialized controllers as operators. This means that operators, such as surgeons, can use their traditional tools to control the robot. In our approach, users interact with an enlarged virtual model of the physical object in virtual reality. Motion retargeting is employed to transmit scaled actions to the robot in a three-stage process: upscaling, constraining, and post-editing the interaction. It is crucial to note that our goal is not to oversimplify surgery. Instead, we aim to use it as a means to demonstrate the potential of our approach to enhance telerobotic collaboration. This involves enabling human operators to leverage the robot's full precision while maintaining the familiar use of traditional tools, rather than relying on specialized controllers. Additionally, we have implemented a preliminary prototype as a foundational step for the ongoing development of the interaction techniques introduced by AntHand.

2 INTERACTION TECHNIQUES FOR TELEROBOTICS

The interaction techniques described by AntHand consist of three phases:

(1) **Up-scaling interaction**: before interaction occurs, the physical model is virtually replicated and up-scaled into the preferred size.
(2) **Constraining interaction**: during the interaction, the effects of movements by the user can be constrained by locking dimensions they are operating in.
(3) **Just-in-time post-editing interaction**: after the interaction, manipulation traces can be further optimized to minimize the noise in the trace or validate for correctness.

2.1 Up-scale interaction

We first address the "up-scale" interaction technique, a concept borrowed from graphical design tools that allow pixel-precise editing operations by zooming in on the manipulated object. It enables users to engage in pixel-level manipulation, increasing precision and accuracy for various tasks such as graphical design, designing electrical components, and medical imaging analysis. We transfer this technique into a virtual environment and tailor it to enable precise manipulation of a robot end effector.

Figure 2 demonstrates the up-scale interaction technique: a seemingly straight line at a lower magnification can, upon up-scaling, reveal itself to be a line with noise. This enhanced visualization afforded by up-scaling is instrumental in facilitating more nuanced and precise interaction, overcoming the limitations inherent in lower-scale representations.

normal scale   zoomed in
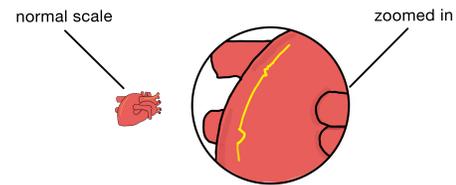
Figure 2: On the left, a standard-scale depiction of a noise-free trace (yellow) is shown. Meanwhile, the right side displays an enlarged view revealing noise within the trace on the heart.

Applying this technique to three-dimensional (3D) interfaces, the "pixel-precise" interaction ensures detailed viewing and precise manipulation of 3D objects. In 3D interfaces, the "pixel-precise" interaction involves enlarging objects in the environment to a desired size. The factor to which extent the object should be scaled, depends on the precision that is expected to be achieved, keeping in mind the constraints introduced by the robotic arm. For example, when a robotic arm has a maximal accuracy of one millimeter, scaling the object beyond millimeter precision, has no particular benefit, as the robotic arm simply cannot get this precise.

However, this approach proves beneficial when precise movements form the basis for interacting with physical models. Such interaction includes surgical scenarios involving anatomical models or assembling compact circuit boards with tiny components.

As seen in previous work [8, 10, 11], diverse methods have been explored to implement the up-scale interaction technique in virtual environments. Commonly used methods typically involve either duplicating a virtual object or adopting a specific viewpoint on objects as a proxy to facilitate interactions. The proxy can be scaled to make interaction easier. Interactions performed on the proxy are applied directly on the virtual object it represents [8, 10, 11]. Similarly to "pixel-precise" interaction, we aim to use the up-scaling technique to enhance the precision and accuracy of object manipulation by offering users a more detailed view of the physical models represented digitally. This behavior is backed by the work of Yu et al. [11], which found that up-scaling models in a virtual environment improved the accuracy of precise annotations. As a side note, our techniques share similarities with the approach presented by Mine et al. [7]. However, a distinctive feature of our methodology is that we apply scaling exclusively to the manipulated objects, departing from the broader scaling of the entire virtual environment proposed in their work.

2.2 Constraining interaction

Interacting with 3D objects in VR using a tracked interaction device, especially when precise manipulations are required, can be cumbersome because of the degrees of freedom and the lack of a physical reference frame. As haptic feedback is challenging and expensive in virtual reality [1], we address this by constraining the interaction of the user. We accomplish this by allowing users to lock the dimensions they are operating in. This means that the user can decide which axis, or combination of axes, they can lock on a specific value. These axes are the x-, y-, and z-axis and the rotational axes of the robotic arm. The techniques, illustrated in Figure 3, contribute to a more intuitive and controlled user experience in VR-based human-robot interaction. This design aligns with Bowman et al.'s guidelines [2], emphasizing the importance of minimizing degrees of freedom in 3D interaction techniques.
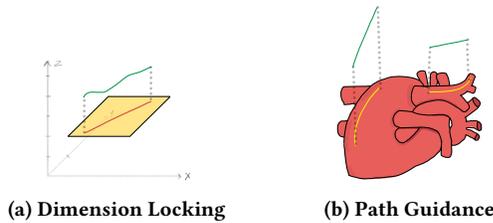

(a) Dimension Locking   (b) Path Guidance

Figure 3: The two constraining interaction techniques proposed in AntHand.

These two techniques, i.e., dimension locking and path guidance, form an inherent advantage of operating within a virtual environment. The main idea behind this is the utilization of motion retargeting, where the actual path is mapped onto the expected path. Figure 3a visually illustrates z-axis dimensional locking. A 3D path (green) is projected onto a plane, thus ignoring z-axis deviations. This method enables users to optimize their positioning or derive feedback by placing a tool on a surface.

In cases where dimension locking proves insufficient, several challenges arise. This scenario is exemplified in Figure 3b, where a complex shape, such as a heart, is used. The expected trace, in yellow, follows the complex shape of the heart. Relying on the "dimension locking" technique will thus fall short, as the complex surface/contour is hard to approximate using ground planes. For this reason, we implemented a "path guidance" technique, which allows the trace in green to be projected on the surface/contour of the 3D model. Implementing such a technique has the advantage that the user does not have to worry about following the surface/contour exactly, and a more suitable or relaxed position can be adopted during the manipulation.

Furthermore, the simultaneous utilization of these techniques manifests their synergy. In the medical context, this could be used for precise straight-line incisions. The path guidance technique would be used to trace the surface/contour of the heart. The dimensional locking technique would be used to lock the y- and z-dimension, enabling one-dimensional tracing. This ensures accurate scalpel positioning, facilitating precise straight-line incisions on the heart.

2.3 Just-in-time post-editing

Figure 4 illustrates the concept of post-editing: after a line tracing task, the user can correct the path to ensure the robot will trace a straight line. Since the user executes all actions in a virtual environment, a delay can be applied before the robot executes the specified actions. Before the path is transferred to the robot, users can still make minor adjustments to provide corrections for potential deviating interactions. These corrections are possible using two approaches: (1) manual correction and (2) digital correction. The manual correction approach relates to manually correcting traces on the model. The digital correction approach utilizes noise filtering algorithms, e.g., Kalman filters, to automatically filter out noise introduced in the trace. After this step, the traces are ready to be communicated back to the robotic arm.
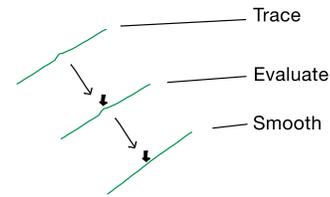

Figure 4: The post-editing process depicted using a three-step example: (1) tracing the line, (2) evaluating noise in the traced line, and (3) smoothing out the noise.

3 ANTHAND IMPLEMENTATION PROTOTYPE

As a starting point for further development of the interaction techniques proposed by AntHand, a basic Virtual Reality (VR) prototype is implemented. This prototype is designed in Unity[1], where direct interaction with the tools known to the operator in VR is supported. The tracking of the interaction tool, the three interaction phases, and the communication back to the robotic arm form the main workflow implemented in the basic prototype. This workflow, whether real-time or in sequence, ensures that precise anticipated virtual behavior translates into the expected physical behavior.
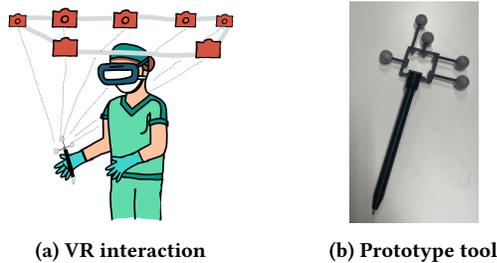
3.1 Tracking manipulation

Accurate tracking of the operators' movement is fundamental, as it is important to define a correspondence between physical and virtual environments [2]. In our prototype for human-robot interaction, we employed an outside-in optical tracking system (OptiTrack[2]) to monitor the movements of a physical tool held by the operator. The use of any tool with attached infrared reflective markers is possible, but it is advisable to utilize tools familiar to the operator, minimizing the need for additional training. Leveraging familiarity with known tools, especially those commonly used by operators, facilitates a seamless integration process. By 3D-printing the tracked tool, replicas can easily be made whilst also keeping the possibility of integrating the reflective markers. On top of this, 3D printing also ensures safety, as the VR headset will prevent the

---

[1]Unity - Real-time development platform
[2]OptiTrack - Flex 13 cameras

user from acknowledging the tool in real life, keeping the user from anticipating sharp edges.

Figure 5b illustrates a 3D-printed pen tool used for interaction in the virtual environment. Five reflective markers are strategically placed on the pen to minimize occlusion between markers while using the tool. Figure 5a depicts the OptiTrack setup used for tracking the pen, held in the surgeon's hand.

(a) VR interaction **(b) Prototype tool**

Figure 5: On the left, an outside-in set-up demonstrates the functioning of the tracking setup, while on the right, a representation displays a 3D-printed pen with five reflective markers attached to it.

3.2 Up-scaling manipulation

The tracked coordinates are transmitted to the Unity[3] prototype, enabling scaled manipulation. The up-scaling algorithm, depicted in Figure 6, involves a user-set scaling factor and calculates the distance between two consecutive coordinates. However, the consecutive coordinates are 3D coordinates, meaning that scaling might become difficult. For this reason, the z-axis calculations are done independently of the x- and y-axis. Figure 6b illustrates the calculations along the z-axis (height). By doing this, the scaled distance along the z-axis can easily be calculated by subtracting the second z-coordinate from the first z-coordinate and multiplying it with the scaling factor. Figure 6a illustrates the calculation along the x- and y-axis. We take both axes independently whilst scaling them but combine them using the Pythagoras theorem to achieve the correct scaled distance.

Operating within an up-scaled environment necessitates downscaling these manipulations before relaying them to the robotic arm. The down-scale process takes on the same approach. However, instead of multiplying with the user-defined scaling factor, the distances between consecutive points now should be divided by the scaling factor. When the computations are performed, the interaction cycle defined in AntHand is complete.

4 DISCUSSION AND CONCLUSION

AntHand introduces a novel approach focusing on precision and accuracy enhancement in telerobotic environments. This approach integrates three distinct interaction techniques: up-scaling, dimension locking, and post-editing, with the primary aim of achieving meticulous manipulation of intricate models. A prototype implementation of AntHand is implemented as a starting point for further development of the interaction techniques.

---

[3]Unity - Real-time development platform

(a) x- and y-axis scaling **(b) z-axis scaling**
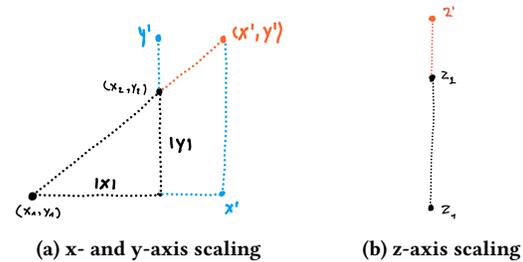
Figure 6: On the left, the resulting point (x', y') is the result of scaling along the x- and y-axis. On the right, the scaling along the z-axis is defined. The red extension depicts the length given by up-scaling the distance between z1 and z2, resulting in z'.

One challenge this approach poses involves accurately replicating complex models, such as a heart, encompassing specific characteristics like blood vessels or rhythmic motion. Neglecting these features during replication might lead to erroneous assumptions during model manipulation, resulting in unpredictable behavior. A potential mitigation strategy involves selectively reconstructing only relevant parts of the model, deviating from fully reconstructing the complex model. However, exploring alternative methodologies, such as video-based interaction akin to the Da Vinci robot[4] or highly precise modeling techniques, is imperative and warrants further exploration.

Another challenge pertains to the robotic arm's speed during movement scaling. Scaling movement inherently alters the arm's actual speed, potentially creating inconsistencies when users expect the arm to mirror the traced speed in the virtual environment. Our proposed interaction cycle, involving a post-editing phase enabling user-controlled arm speed, overcomes this issue. However, speed-related challenges may arise in scenarios requiring real-time interaction without the post-editing phase.

In addition to the proposed challenges, future work in AntHand involves (1) studying and integrating interaction techniques used in existing computer-aided design (CAD) software, such as Fusion 360[5], Blender[6], and Meshmixer[7], (2) initiating the interaction techniques, and (3) evaluating the effectiveness of AntHand in use cases that demand precise manipulations.

Despite the outlined challenges, the ongoing development and refinement of AntHand showcase its potential to significantly advance the precision and accuracy of telerobotic systems. Future investigations into these critical areas are pivotal to further enhancing the capabilities and applicability in precise environments.

ACKNOWLEDGMENTS

---

[4]Intuitive - da Vinci
[5]Autodesk - Fusion 360
[6]Blender - 3D software
[7]Autodesk - Meshmixer

REFERENCES

[1] Dr Bishop, Reader Dr, Anselmo Lastra, and Mark Mine. 2001. Exploiting Proprioception in Virtual-Environment Interaction. (May 2001).

[2] Doug Bowman, Ernst Kruijff, Joseph J. LaViola Jr, and Ivan P. Poupyrev. 2004. *3D User Interfaces: Theory and Practice*. Addison-Wesley. Google-Books-ID: JYzmCkf7yNcC.

[3] Chadrick R. Evans, Melissa G. Medina, and Anthony Michael Dwyer. 2018. Telemedicine and telerobotics: from science fiction to reality. *Updates in Surgery* 70, 3 (Sept. 2018), 357–362. https://doi.org/10.1007/s13304-018-0574-9

[4] A. Gasparetto and L. Scalera. 2019. A Brief History of Industrial Robotics in the 20th Century. *Advances in Historical Studies* 08, 01 (2019), 24–35. https://doi.org/10.4236/ahs.2019.81002

[5] C Heyer. 2010. Human-robot interaction and future industrial robotics applications. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, Taipei, 4749–4754. https://doi.org/10.1109/IROS.2010.5651294

[6] R. Judd and Al. Knasinski. 1987. A technique to calibrate industrial robots with experimental verification. In *Proceedings. 1987 IEEE International Conference on Robotics and Automation*, Vol. 4. Institute of Electrical and Electronics Engineers, Raleigh, NC, USA, 351–357. https://doi.org/10.1109/ROBOT.1987.1088010

[7] Mark R. Mine, Frederick P. Brooks, and Carlo H. Sequin. 1997. Moving objects in space: exploiting proprioception in virtual-environment interaction.

In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques - SIGGRAPH '97*. ACM Press, Not Known, 19–26. https://doi.org/10.1145/258734.258747

[8] Jeffrey S. Pierce, Brian C. Stearns, and Randy Pausch. 1999. Voodoo dolls: seamless interaction at multiple scales in virtual environments. In *Proceedings of the 1999 symposium on Interactive 3D graphics*. ACM, Atlanta Georgia USA, 141–145. https://doi.org/10.1145/300523.300540

[9] Christian Weckenborg, Karsten Kieckhäfer, Christoph Müller, Martin Grunewald, and Thomas S. Spengler. 2020. Balancing of assembly lines with collaborative robots. *Business Research* 13, 1 (April 2020), 93–132. https://doi.org/10.1007/s40685-019-0101-y

[10] Jacob Young, Nadia Pantidi, and Matthew Wood. 2023. I Can't See That! Considering the Readability of Small Objects in Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics* 29, 5 (May 2023), 2567–2574. https://doi.org/10.1109/TVCG.2023.3247468

[11] Kevin Yu, Alexander Winkler, Frieder Pankratz, Marc Lazarovici, Dirk Wilhelm, Ulrich Eck, Daniel Roth, and Nassir Navab. 2021. Magnoramas: Magnifying Dioramas for Precise Annotations in Asymmetric 3D Teleconsultation. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, Lisboa, Portugal, 392–401. https://doi.org/10.1109/VR50410.2021.00062