Team VolksDroid Avatar

The VolksDroid team is creating a Remote Reality™ (RR) telerobotics experience that is immersively high quality, inventive, and economical. Our goal is to create a service providing useful, inexpensive avatars and operator rigs for use in elder care, home service, disabled proxies, and beyond.

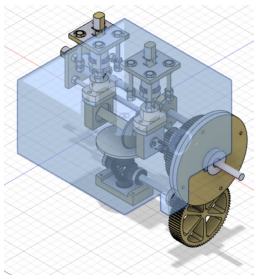
The primary measures of success are fidelity of experience of the operator, dexterity and capability of the robot, economy and manufacturability, and strength of system architecture supporting open innovation, security, and extension. We expect to make major improvements in core robotic and avatar enabling subsystems of actuators, sensor fusion, complex touch and haptics, and AI / ML safety, intent, tasking, and an overall open extensible system.

Avatar Form

The VolksDroid team will start with a humanoid robot with a wheeled base, later switching to legs. The main rationale for a humanoid form is that an operator will most easily inhabit a humanoid form.

We have a new patent-pending actuator solution, the SPIVT (Slip-modulated Prioperceptive Infinitely Variable Transmission), that could drastically lower the cost, weight, and complexity of robotic actuators. Options include the Innfos / Cloudminds XR-1 or Modbot 7DOF arm systems while we develop. Innfos and Modbot have compact gearmotors suitable for direct actuation or for driving SPIVT RIVT (reversible infinitely variable transmission) actuators.

Innfos SCA

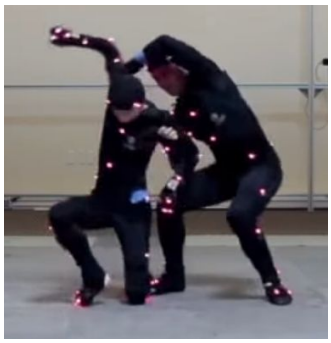


SPIVT

Starting with a customizable version of available hand models, we will add our own high-resolution touch sensor. The robot head will be covered with OLED panels, showing the current operator's eyes, mouth.

Operator Rig

A commercial haptics suit, the Holosuit, will be used as the operator rig starting point. VolksDroid team will test alternatives like PhaseSpace w/ goggles (expensive) and Oculus Quest (inexpensive, more mobile) for accurate positional tracking and RR goggles. Remote Reality™ makes use of AR/VR methods coupled with full sensory situational awareness of the avatar and its environment. This includes accomplishing SLAM of the environment. Initial SLAM in a room will use a floor-roaming development unit that can provide SLAM results, shared via Wi-Fi.



PhaseSpace

Will test a variety of haptics gloves, such as Holosuit and VRFree gloves. We will also construct our own haptics gloves: low-profile, wearable gloves using sheathed cables attached to fingers, pulled precisely by custom voice coils. This provides direct and transverse pressure with some resolution while allowing hands to look normal to hand tracking in the Oculus Quest or similar.

Sensory Modalities

The avatar vision system will consist of several sensors combined to provide, to both the operator and AI/ML assistance subsystems, a complete and high-quality sense of the avatar situation. This

situational awareness contributes to the sense of immersiveness for the operator, eliminates extra looking around, and is essential for automated safety and task planning and execution.

The primary sensors for vision are the main high-resolution cameras, including cameras at different focal lengths for zooming, plus initially, the DreamVu stereo depth camera. Additional 360-degree cameras may be used individually or in stereo, such as the Ricoh Theta V. Other sensors include: proximity sensing IR emitter/sensors, low (64 pixel) and high-resolution time of flight cameras, IR/FLIR sensors, and radar systems.

Positional sound provides subtle signals for the presence and direction of people, machines, and orientation in a room. The VolksDroid avatar will have an array of microphones, recording from every direction. This allows the operator's perception of sound to track avatar and/or operator head position in the soundspace.

There are multiple uses of haptics in the VolksDroid avatar: Delivering the sense of touch, haptics for body position, situational awareness, and other channels of information such as compass / room orientation.

The remote sense of touch is the most important haptics need. The avatar hands and some part of the body will be covered with high-resolution touch sensing. This includes direct and transverse pressure, vibration and texture, and heat dissipation indicating material type. The operator rig will transmit the results of this sensing to the operator.

Our prototype touch sensor is a skin of silicone rubber, optical fibers, and a modulator / detector that minimize connections and electronic components. Haptic sensing will involve custom voice coils actuating cables to fingertips. Also, modulated electrostim will be explored.

AI, ML, Latency, Safety, Operator Experience

AI/ML: on-board, OCR, voice recognition, Ortelio cloud services.

Real-time processing on the avatar and the rig, low-overhead protocols minimize intrinsic latency.

Intrinsic safety provided to prevent mishaps: Sensor fusion for situational awareness, proximity sensors such as IR emitter/sensor pairs, and task feedback loops such as auto-pause.

VolksDroid will have a task planning, pre-simulation system: an operator identifies tasks for the avatar to accomplish semi-autonomously, such as moving to something, picking up an object, opening the refrigerator door. The operator can then work ahead or temporarily disconnect their pose from that of the avatar.

Primary Objectives

Objective 1 - Touch + Haptics

High-quality touch experience requires both high-resolution robotic touch sensing and corresponding operator haptics. Our solution is a thin, resilient layer that can be laminated on robot surfaces requiring few connections to a microcontroller for sensing. This uses silicone rubber, treated optical fibers for selectively sending and receiving light, a sender unit, and a receiver unit, along with a microcontroller.

Objective 2 - Sensor Fusion

Sensor fusion provides a forward image plus a surround strip. Image understanding includes face, person, poise detection with YOLO and OpenCV locally, plus Ortelio cloud processing. Operator can zoom part or all of the screen; later, eye tracking.

An important HRI capability is synchronization of avatar body position and operator current position. We will be testing strategies and algorithms for helping an operator get in sync, but also for quickly connecting and suspending pose tracking to allow breaks, interruptions, and to let semi-automated tasks to complete independently.

Objective 3 - RIVT Actuators - SPIVT

Existing robotic solutions are expensive, heavy, inefficient, and usually clunky. By using a new combination of mechanisms, we can drastically lower cost while also lowering weight, getting a variable gear ratio, and often improving form factor.