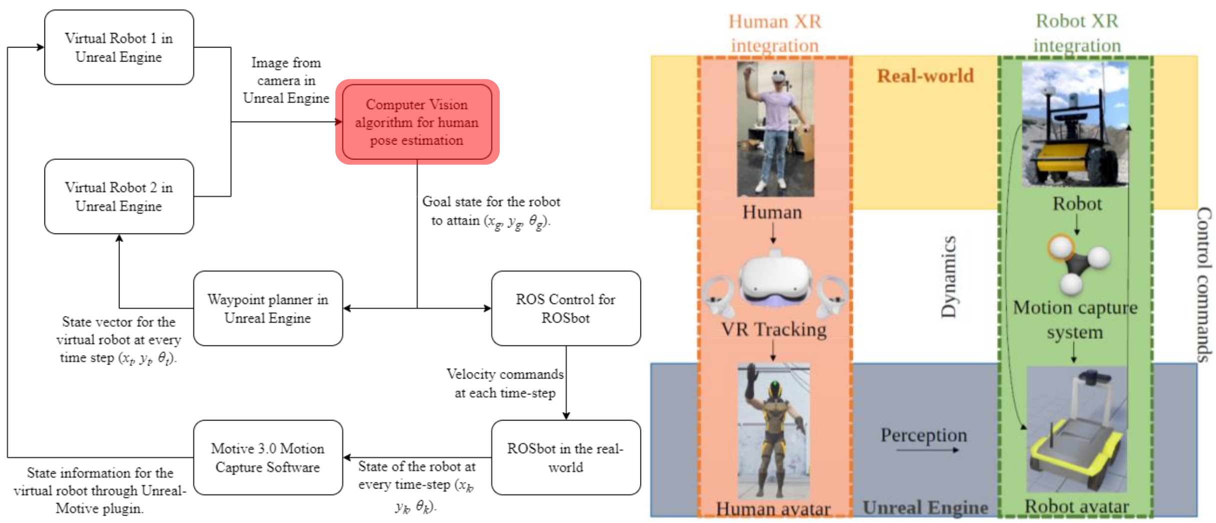
CS6670 Computer Vision: Project Final Result

Human pose estimation for robot control in a cyber-physical virtual reality testbed for human-autonomous systems collaborations

The flowchart of the architecture that we designed and developed for our system to enable human robot collaborations in a cyber physical environment is shown in Figure 1 (a). Computer Vision forms a vital component of this facility. It is through computer vision algorithms that the robots (both physical and virtual) present in the testbed are able to interact with the individual operating the facility. For this course project, we implement human pose estimation using the state-of-the-art real-time, multi-person keypoint detection library: OpenPose. Figure 1 (b) shows how we integrated the human into the cyber-physical environment using Virtual Reality headset and controllers. It also shows how we created a digital twin for the real robot (a Jackal robot in our case) into our environment using the OptiTrack motion capture system.



(a) (b)

Figure 1: Architecture of the virtual reality testbed. (a) Flowchart explaining the system's architecture. OpenPose serves as the primary computer vision algorithm. (b) Pictorial representation of the cyber-physical integration of the human operator and the robot in the facility.

Control commands are passed to the purely virtual & cyber-physical robot using OpenPose. The operator uses different poses which provide control commands to the robots. For this project we use 3 poses (shown in Figure 2) to provide rightward motion (a), forward motion (b), and leftward motion (c) to the robots present in the facility.

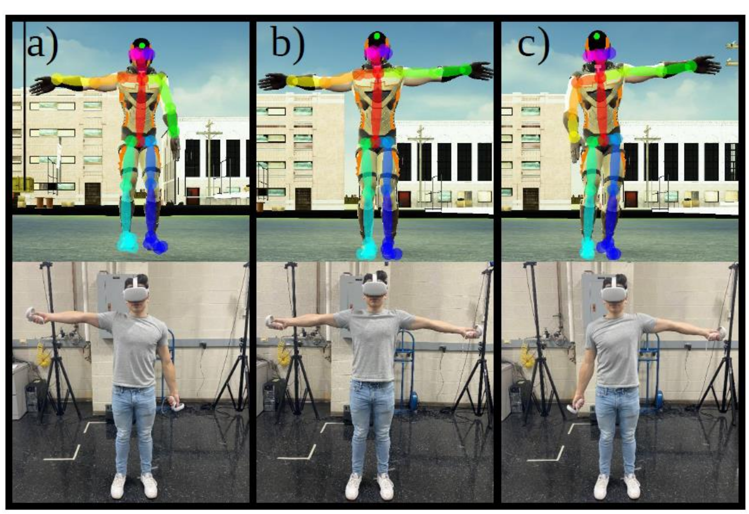


Figure 2: OpenPose detection of the human avatar in the virtual environment for providing control commands.

Figure 3 shows the results obtained for one run. The trajectories of the purely virtual, cyber-physical, and purely real robot are shown. We use the 3 aforementioned poses to provide 17 control commands to the robots in a single run. The real robot and its cyber-physical twin start 1 meter behind the purely virtual robot.

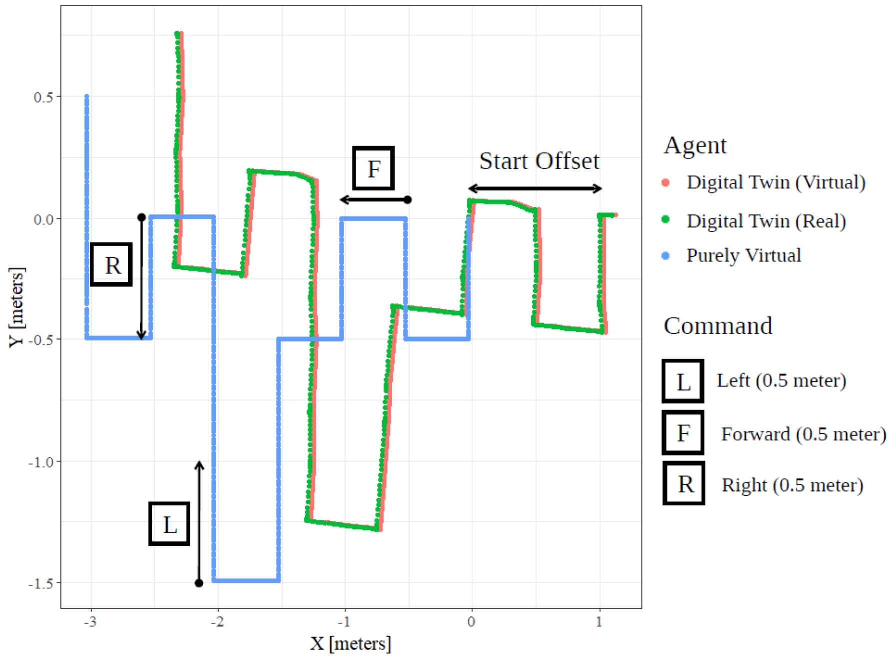


Figure 3: Results for a single run in the testbed.

As expected, the purely virtual robot moves exactly as instructed, however, the real robot and its cyber-physical counterpart are not able to move with the same accuracy & precision. The reason being twofold: i) the major problem being the biases & limitations on the accuracy of the motion capture system, ii) the minor issue with drift & friction experienced by the real-world robot. Nonetheless, both robots followed the control commands which indicates that the architecture we developed for the testbed functions as per our requirement. This result demonstrates the feasibility of our facility to be used as a real-time human autonomous systems collaboration environment. While we use human pose estimation as the computer vision algorithm for this course project, it is important to note that the facility can be used to evaluate the performance of any such algorithm in different environmental settings on robotic platforms.

This facility provides us best of both worlds, i.e., virtual perception which allows us to test various computer vision algorithms in any setting that we desire (not being limited by the physical environment) and accurate dynamic behavior of different robots being captured by the motion capture system (thus avoiding modeling & linearization of highly nonlinear & coupled robot dynamics). Our further work will focus on the following:

1. Account for the biases and limitations of the motion capture system.
2. Implement multi-agent (robot) collaborations (such as consensus control) with the commands being given by the human in the facility.
3. Evaluate recognition-based algorithms using this facility on different robots.