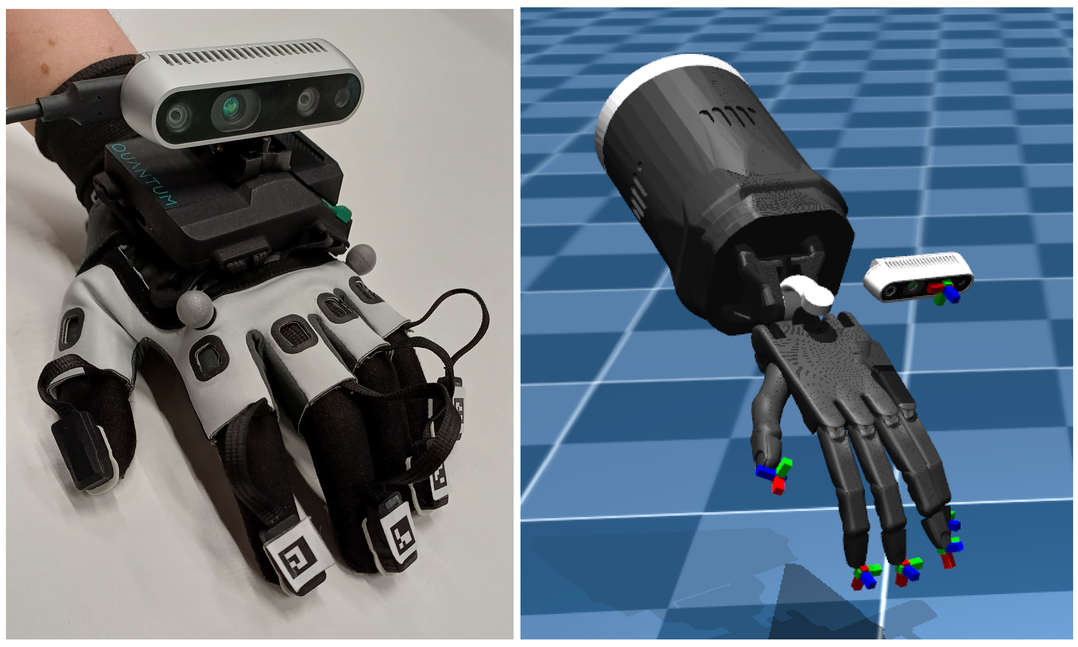
Towards Dexterous Robotic Manipulation in Virtual Reality

Theo Krijgsheld

Master's Thesis Project. Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen

Abstract

Dexterous robotic manipulation is a crucial research field in robotics and embodied AI. However, most of the data collection systems rely on complex user interfaces in reality. In this research, we worked towards a more intuitive method to control dexterous robotic hands in virtual reality using finger motion capture gloves in combination with visual-inertial odometry for absolute pose tracking. This allowed us to move the virtual robotic hand in three-dimensional space relative to a defined world frame. During our research we experimented with tasks like grasping and placing objects, but we also experimented with two handed interactions, like passing an object from one hand to the other. While we were able to execute the tasks using the intuitive control over the hands, the simulation showed the hands in a pose that did not accurately match the pose of the real life hand. This issue is caused by a scaling issue and drift in the visual-inertial odometry and could not be solved during this project. Therefore, future work might include adjusting the visual-inertial odometry to get precise poses more often.



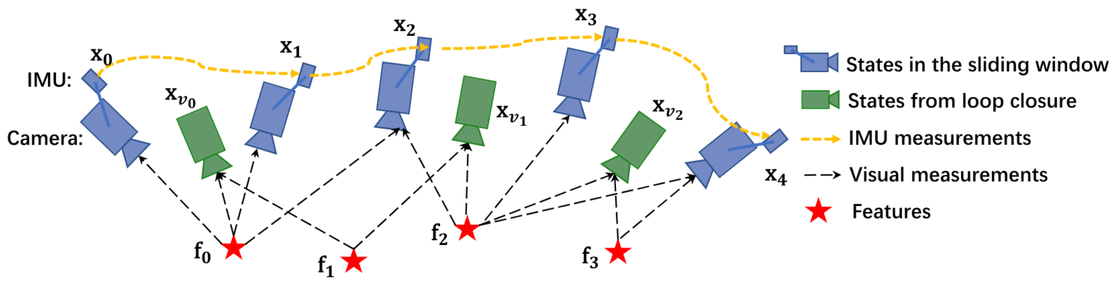
The glove in real life and the virtual hand in MuJoCo

Background

Dexterous robotic manipulation refers to a robotic hand performing complex tasks involving manipulation of objects. It is a crucial research field already applied in industry, for holding tools, surgery, for robot assisted training, and in space and underwater exploration. However, so far most robots use either a simple two fingered gripper or a complex user interface which causes an unintuitive control over the hand. Using finger motion capture gloves, we aim to create an intuitive way to control a 5-fingered humanoid robotic hand in virtual reality.

Methodology

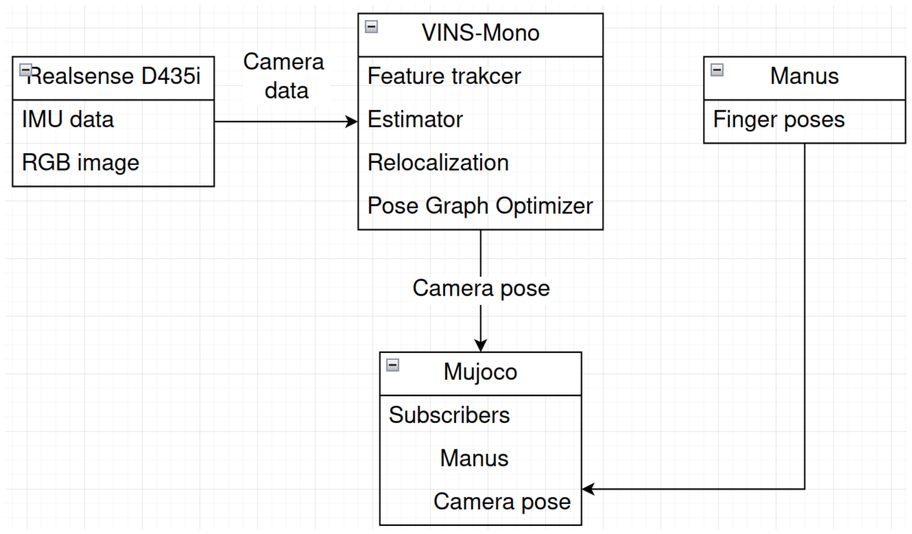
The Manus Quantum Metagloves use an electromagnetic field to measure the pose of the fingers relative to the hand. The data can be extracted using the Manus SDK and it can be used to control a virtual hand. The pose of the camera we attached to the glove can be estimated using a Visual Inertial Odometry (VIO) system. For this we used VINS-Mono, which is an optimization based VIO, including a relocalization process to eliminate drift and a global pose graph optimization to ensure global consistency.



An illustration of VINS-Mono

We combined the Manus gloves with cameras and simulated the virtual hands at the pose that was estimated by VINS-Mono. We also used the Manus Gloves ROS2 driver to set the finger poses relative to the hand pose. Furthermore, we calibrated the transformation between the camera and the base of the glove using ArUco markers on the fingertips. The markers can be detected using OpenCV, giving a transformation between the camera and the markers. The glove measures the transformation from its base to the fingertips. Using this information, we can also compute the transformation between the camera and the base of the glove. We used an ArUco marker to define the world frame as well.

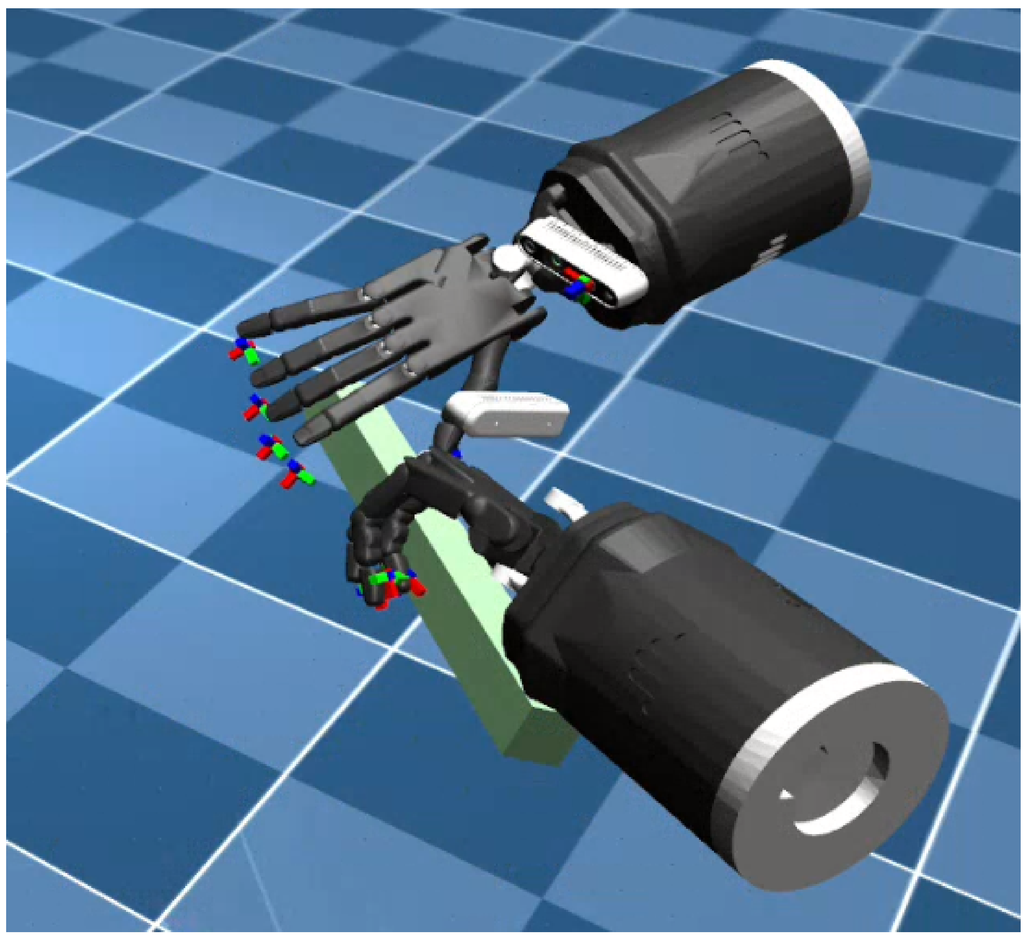
The pipeline of our system is depicted below. Furthermore, the calibration process using the ArUco markers is executed before VINS-Mono starts. The data it collects is forwarded to the MuJoCo simulation, which uses it to set a relative pose between the camera and the virtual hand. Additionally, the MuJoCo client transforms every incoming camera pose using the data from the world frame calibration.



The pipeline of our system

Evaluation & Conclusion

With the described setup we were able to fully control both the three-dimensional pose of the hand and the poses of the fingers. We were able to do tasks like picking up objects, placing them back down and passing an object between two hands. However, we also tried grasping objects that were modeled after real life. As we were not able to do so, we concluded that the pose of the virtual hand does not correspond to the pose of the glove in real life. This is likely due to drift from VINS-Mono, or a scaling issue, for which we did not have time to solve it during our project.



Passing an object between two hands in virtual reality

Acknowledgments:

I would like to thank Dr. Kailai Li and Qihao Yuan for their supervision during the project. Their guidance was incredibly valuable in this research. This project is supported in part by the Comenius Teaching Fellows grant under grant number OFUKV18739.

university of groningen

Agile Sensing & Intelligence Group ASIG