

# Roboy Avatar Context Awareness through ToF

## Project Description

The project aims to increase contextual awareness of the operator in the telepresence scenario. During immersive teleoperation it is desirable to provide the operator with the high fidelity representation of the remote environment perceived by their robotic avatar. In the current setup, this is partially achieved through stereovision. The field of view of the cameras used, however, is smaller compared to the FoV of the HMD (assuming real-life scaling) and significantly smaller compared to the FoV of the human eye. Additionally, since the current camera setup does not support saccades (which play an important role for human vision), the stereo-effects are distorted.

This project aims to integrate a long-range ToF camera into the robotic avatar system to increase the fidelity of the remote environment reconstruction in 3D. Through combining the depth and RGB image streams and reprojecting them into the 3D space, the robot's environment is to be reconstructed in real-time. This data together with the camera pose estimate is then to be streamed to the operator's software where a virtual 3D scene is to be created. This scene is subsequently to be fused with the stereo-camera stream to generate an immersive first person view perspective displayed in the HMD.

## System Requirements

- Mid-range immediate environment reconstructions (10 m radius)
- Minimal bandwidth use
- Differential streaming -> only new points
- Computation time below 10ms
- Potentially cloud processing and only streaming of final mesh to headset

## Interface to Roboy

Current avatar on-board computational resources:

- Nvidia Jetson AGX Xavier
- Terasic DE-10 Nano SoC FPGA Dev-Kit
- ROS 1 middleware

Current avatar - operator data transmission stack:

- Animus SDK
- 200 ms latency on 720p camera stream

Operator software stack:

- Oculus Quest 2 (used in both modes: standalone Android and Oculus Link)
- Unity 3D



## Previous Work

We are currently supervising a M.Sc. thesis "Streaming of a real-time dense surface reconstruction to improve spatial awareness in telepresence". The developed system uses Intel Realsense L515 camera and is based on InfiniTAM - an open source, multi-platform framework for real-time, large-scale depth fusion and tracking. It is fully integrated into Roboy's telepresence stack. User studies were performed and the resulting data is currently analyzed.

Developed system shortcomings:

- Small range (2m radius)
- No localization
- Reconstruction is not performed in real-time

Given the progress of development and the constraints this system exhibits, we envision that a de novo development based on Infineon and PMD's work is better suited (replaces Docker Container A in the image below). The integration into our Animus transport stack and the visualisation on our Unity client, can be reused.



