



MOTIUS
WE R&D.

Crane Tele-Operation

Motius GmbH

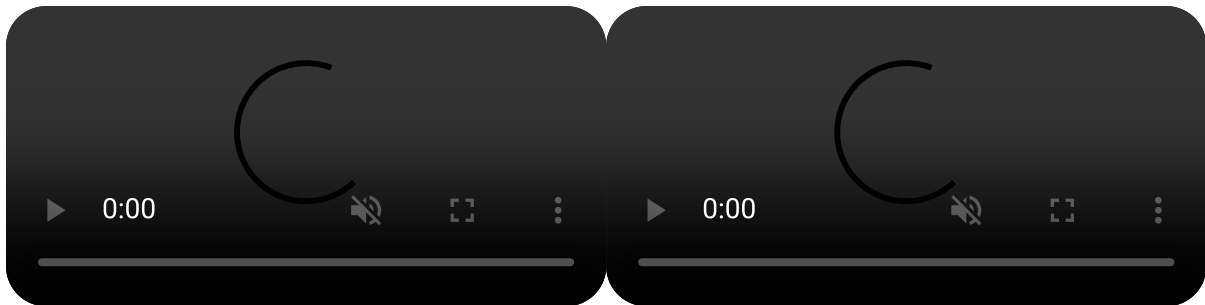
January 08, 2026 17:42 (4e53b49)



Crane Tele-Operation

Heavy Machinery

Computer Vision



For Sennebogen, we built a teleoperation system that was shown on bauma 2025. It features a virtual shadow depth visualization, while streaming video & control signals with low latency (sub 100ms, including AI inference for depth visualization and object classification).

Shadow Projection

A key part of the user interface is a virtual shadow of the implement that gives the user a way to "aim" in all three axis:



Depth projection and control panel

1. The stereo camera creates a point cloud for the current frame
2. The point cloud is rotated by the camera IMU vector so it is aligned with the gravity vector
3. We transform the last known implement position from CAN into this point cloud coordinate system
4. Highlight pixels in the image that are near the implement position, based on their 3D point cloud position

Hardware

The key components of this system are:

Component	Description	Price
Stereolabs ZEX X AI camera	High-performance camera for depth perception and 3D vision with integrated IMU	~\$400
NVIDIA Jetson Orin	AI computing platform for edge inference and real-time video processing	~\$2,000
PCAN-Ethernet Gateway	CAN bus gateway for vehicle communication	~\$400
Teltonika RUTX50 with Tailscale	Industrial 5G router with VPN for secure, low-latency remote connectivity	~\$500

Software

On the **sender** side two cameras are connected to the Jetson Orin via GMSL. The Stereolabs SDK renders a point cloud from the calibrated pair of ZED X One cameras.

The **receiver** opens a connection to the sender and decodes the stream of point-cloud & video data, using a modified version of ffmpeg with improved hardware encoding & decoding.

The receiver runs the user interface, which decodes the stream, and connects to the PCAN gateway to receive important status messages, as well as the implement coordinates.

Network connectivity between sender & receiver is provided by a set of Teltonika routers, which contain a failover to a secondary Starlink network if the primary connection fails. The backup network introduces a lot of additional latency, but still offers some limited functionality for tele-operation. Video stream & control signals are encrypted through a Wireguard VPN, which provided the lowest latency in our testing, compared to other VPN technologies.

The controllers on the receiver side and the machine connect to each other through two paired PCAN gateways.

Key Learnings

Open-Source

Open-source software from Stereolabs & NVIDIA was key to debug & improve the system during development. Proprietary video streaming & camera options were not as fast, or as flexible, and offered no easy way for adaptation.

Autonomous Driving Tech

High-performance sensors & compute modules from autonomous driving and robotics applications have opened up a lot of options for tele-operation

Environmental Challenges

Low light and bad weather reduce performance, especially of depth perception. Camera position (for example on the boom) can help. Lidar & radar are more robust, but still costly