

MANTRA: Mixed-reality Augmented Networks for Teleoperated Robotics Applications

Introduction

The MANTRA project develops a Mixed Reality (MR) interface for teleoperating industrial mobile robots, integrating multimodal feedback—like haptic interfaces—to improve user experience and precision. A key element is participatory design, involving stakeholders through co-design sessions to build a robust, user-friendly system. To support real-time bi-directional communication, MANTRA implements software mechanisms that ensure 5G/Wi-Fi interoperability. It also incorporates edge process offloading to reduce mobile robot computation.



Fig. 1: User Interface

Objectives

- Reduce workplace accidents in manufacturing by using MR-enabled telepresence tools.
- Improvement of 5G and Wi-Fi interoperability to ensure Quality of Experience (QoE) in challenging radio environments.
- Enabling XR robot control via edge computing to reduce latency and offload sensor data processing.

Architecture

The MANTRA project features a modular, layered architecture to support safe and efficient teleoperation of mobile robots using MR. It consists of three main components:

MR Interface: Developed for HoloLens 2, the MR application enables remote control of Autonomous Mobile Robots (AMRs).

References

n.d. Accessed 07 04, 2025. <https://github.com/bluenvirion/mediamtx>.

Meseguer Valenzuela, Andrés, Javier Silvestre-Blanes, Víctor Miguel Sempere-Paya, and Luis Miguel Bartolín-Arnau. 2024. "Multi-Access Edge Computing performance into Non-Public 5G Networks: A robot-based experiment." 2024 IEEE 20th International Conference on Factory Communication Systems (WFCS). Toulouse.

Soler-Dominguez, Jose Luis, Samuel Navas Medrano, Marta García Ballesteros, and Patricia Pons. 2025. Design guidelines learned from the MANTRA project. Instituto Tecnológico de Informática, Zenodo. doi:10.5281/zenodo.15753852.

Co-designed with stakeholders, it provides visual, haptic, and audio feedback and offers three control modes: arrow panel, slider, and joystick. Users can switch between RGB, thermal, and point cloud views in real time, depending on task needs. The system integrates with ROS, subscribing to topics like /pointcloud and /battery for live data.

Edge Multimedia Processing: To reduce robot-side computation and latency, sensor data (RGB, thermal, point cloud, battery) is offloaded to edge servers. Using ROS and GStreamer, data is processed and streamed via RTSP pipelines built with MediaMTX and deployed in Docker. A Python ROS node overlays key information on video and republishes it. The system has been tested in SPIRIT's Trusted Execution Environment for secure deployment.

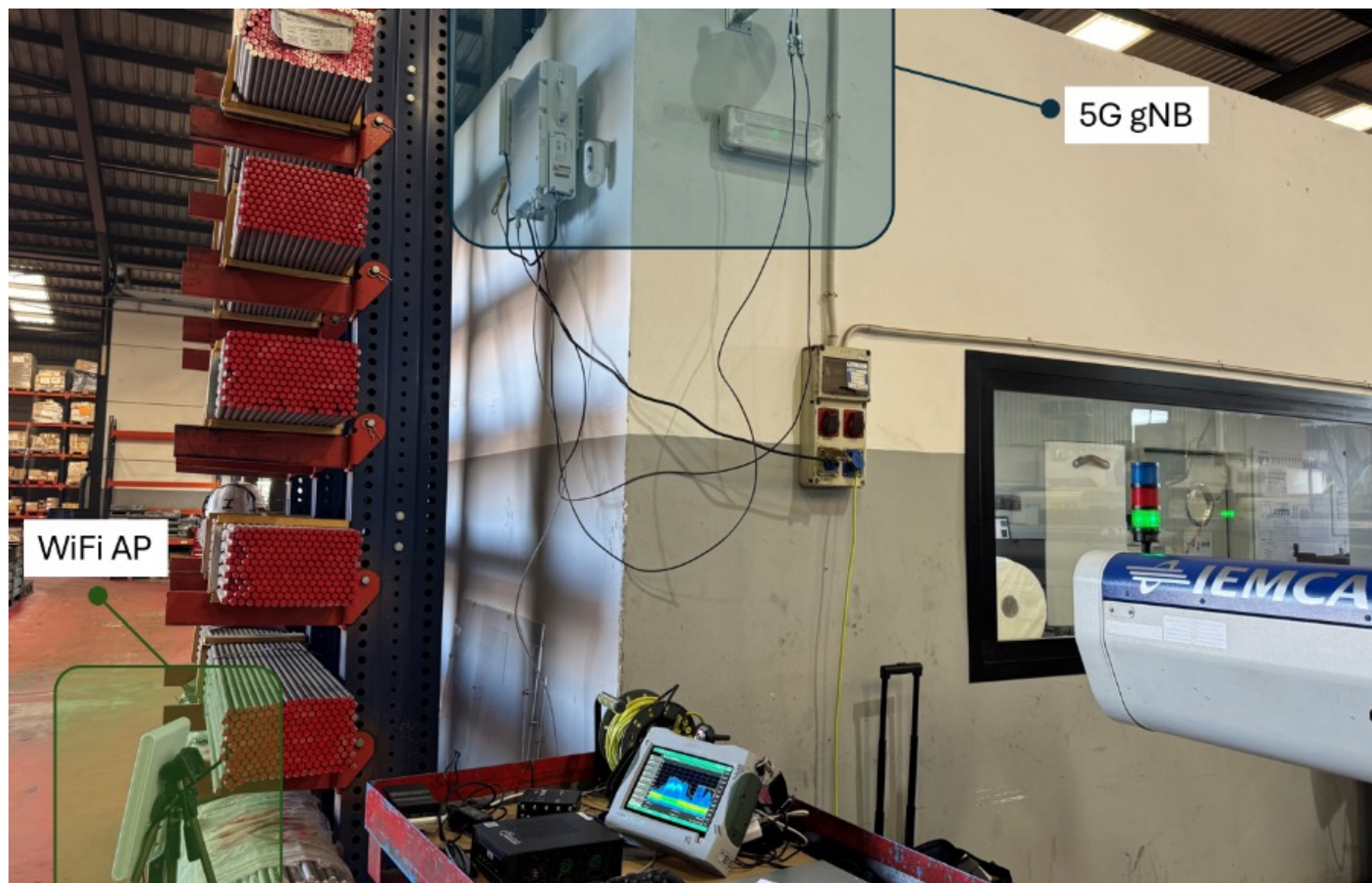


Fig. 2: 5G & Wi-Fi Network deployment

5G–Wi-Fi Interoperability: MANTRA ensures stable robot connectivity using 3GPP's ATSSS to switch between 5G and Wi-Fi networks based on real-time metrics like RSSI and RSRP. A hysteresis mechanism avoids constant switching near signal thresholds, ensuring seamless communication and QoE.

Results

The MANTRA MR teleoperation system was evaluated through pilot and formal usability studies with FACTOR employees and a smaller group at the Berlin testbed. During initial testing at ITI, FACTOR staff explored the MR interface, operated the robot, and gave early feedback.

A later study at FACTOR involved 10 users (ages 23–57) with varying XR experience. Participants used the MR interface to control the robot and completed three standardized questionnaires: USEQ (satisfaction), NASA-TLX (workload), and TPI (presence).

Results show the system offers a usable and immersive experience with moderate workload and high acceptance. Areas for improvement include audio realism and comfort. Connectivity goals for 5G/Wi-Fi interoperability and edge offloading were met, confirming system performance.

- **USEQ** averaged 24.5/30 (>80%), showing strong satisfaction.
- **NASA-TLX** indicated low frustration, effort, and physical demand, with moderate temporal and mental demand (~2.5/5), and best scores in performance (3/5).

Description	Target Value	Result
Users tested	>10 users	13 users
User acceptance	>80% of user acceptance	81.6%
Multimodal feedback	3 sources	haptic, visual and audio
Design guidelines	>= 5	5
5G/Wi-Fi timeout	40 ms	12ms (average)
Video Streams	>=4 x (720p @30fps)	4 x (720p @ 30fps)
Sensor Streams	>=3	Point cloud, thermal and battery sensor

Tab. 1: Results summary

Authors:

Andrés Meseguer Valenzuela (Instituto Tecnológico de Informática, ITI),

ameseguer@iti.es.

Patricia Pons Tomás (ITI)

Samuel Navas Medrano (ITI)

José Vera Pérez (ITI)

Jose Luis Soler-Domínguez (ITI)

Marta García Ballesteros (ITI)

David Millán Escrivá (ITI)

Pablo Rocamora Montoro (FACTOR S.L.)

Alfredo Giménez Millán (FACTOR S.L.)

