The NimbRo Immersive Telepresence System Winning the ANA Avatar XPRIZE Finals

Max Schwarz and Sven Behnke

University of Bonn
Computer Science Institute VI
Autonomous Intelligent Systems

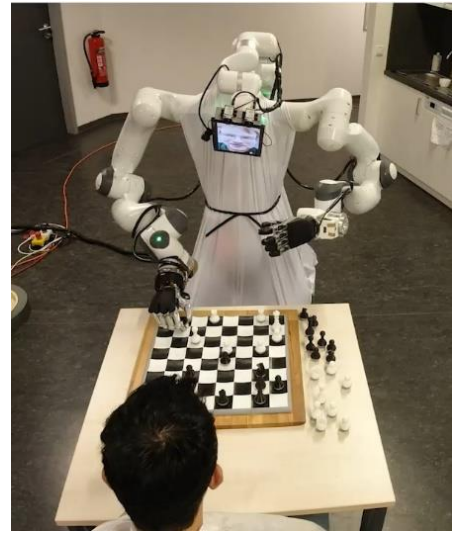
Experience with Teleoperated Robots

- Multiple domains
- Often motivated by competitions and challenges

RoboCup@Home

DARPA Robotics Challenge
DLR SpaceBot Cup

CENTAURO



ANA Avatar XPRIZE

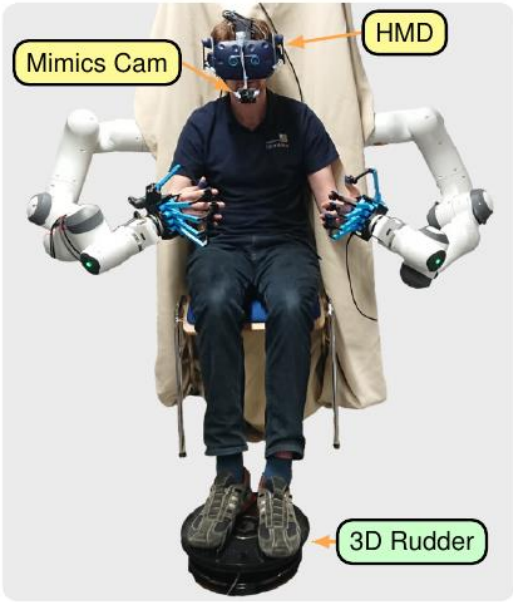
ANA Avatar XPRIZE Competition

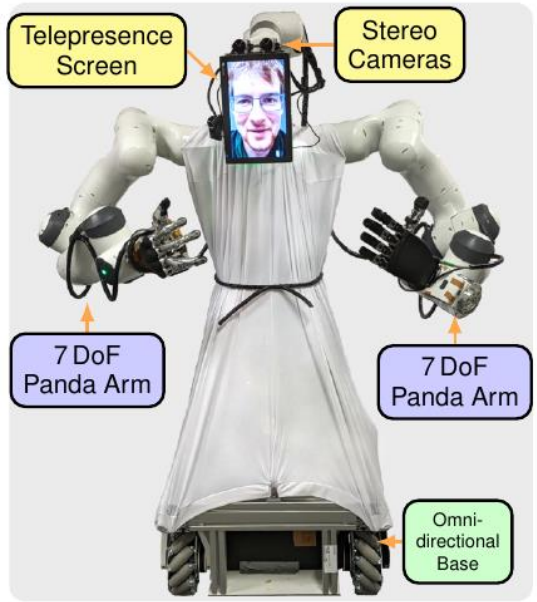
- Requires mobility, manipulation, human-human interaction
- Focuses on the immersion in the remote environment and the presence of the remote operator

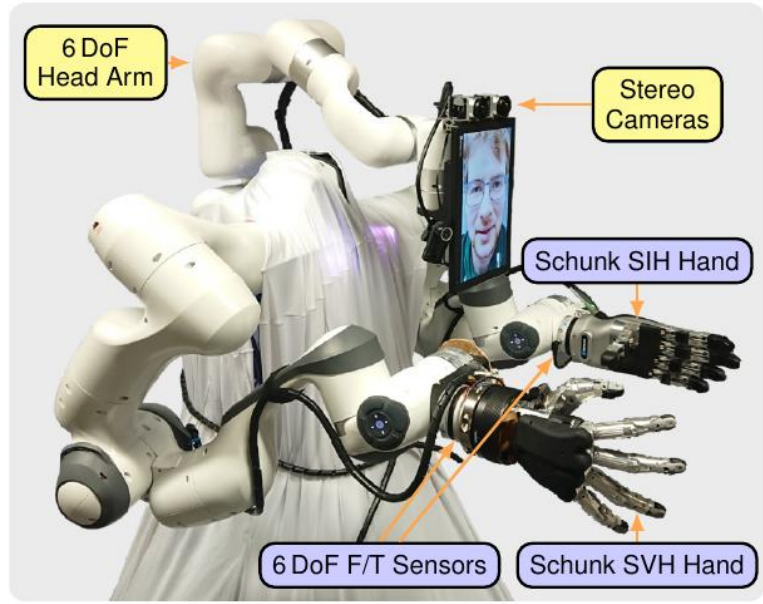
NimbRo Avatar 2021

ANA AVATAR XPRIZE

- Two-armed avatar robot designed for teleoperation with immersive visualization & force feedback
- Operator station with HMD, exoskeleton and locomotion interface





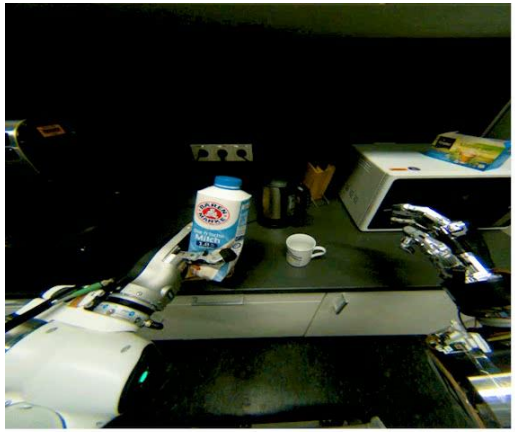


[Schwarz et al. IROS 2021]

UNIVERSITÄT BONN AIS

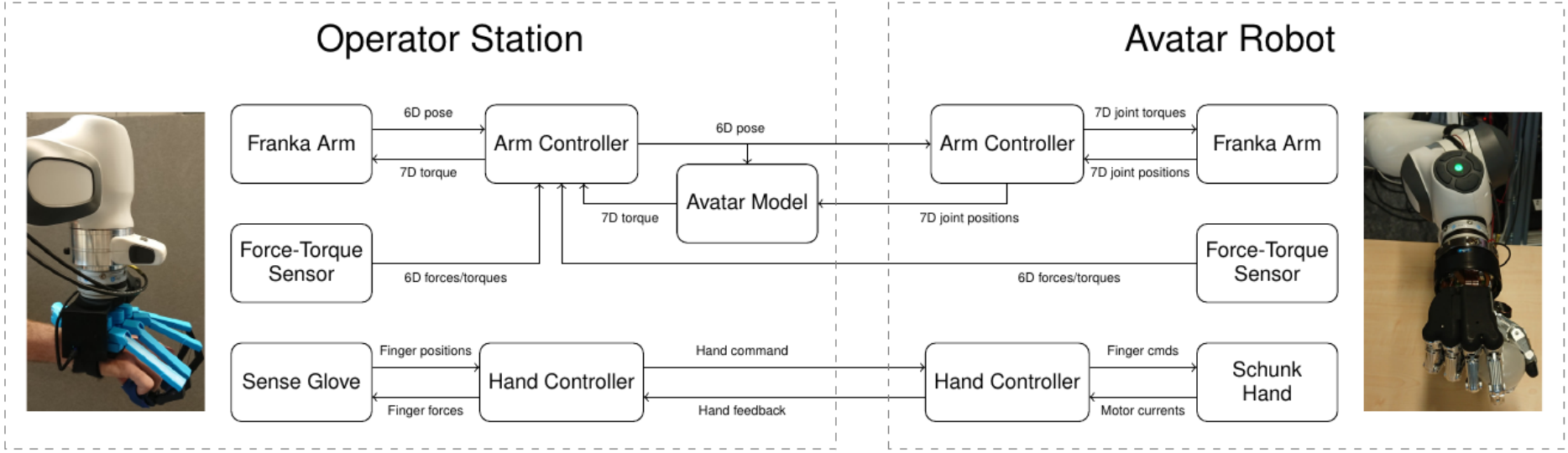
Team NimbRo Semifinal Submission

ANA AVATAR XPRIZE



[Schwarz et al. IROS 2021]

Manipulation with Force and Haptic Feedback



- Arm exoskeleton (Franka Emika Panda), F/T sensor (Nordbo + OnRobot HEX), hand exoskeleton (SenseGlove)
- Avatar side: Arm + F/T sensor + Schunk SVH / SIH hand
- Provides force feedback for wrist and haptic feedback for fingers
- Avatar limit avoidance using predictive model to reduce latencies

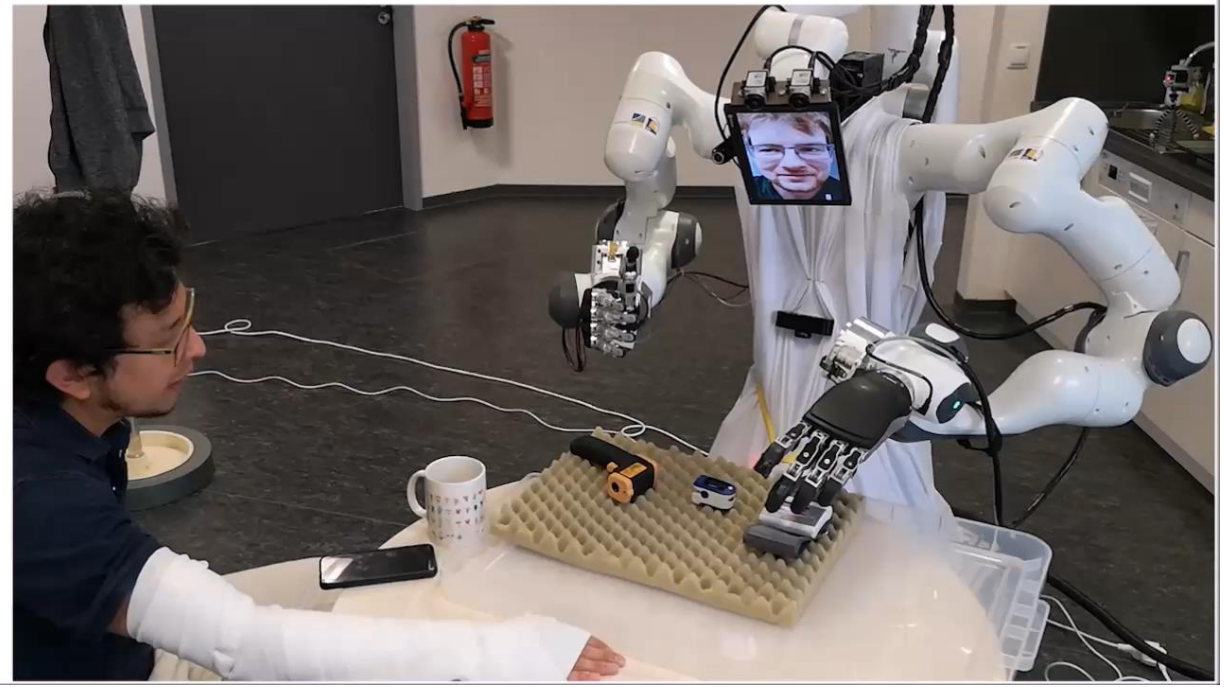
[Lenz and Behnke ECMR 2021]

Team NimbRo Semifinal Team Video

ANA AVATAR XPRIZE

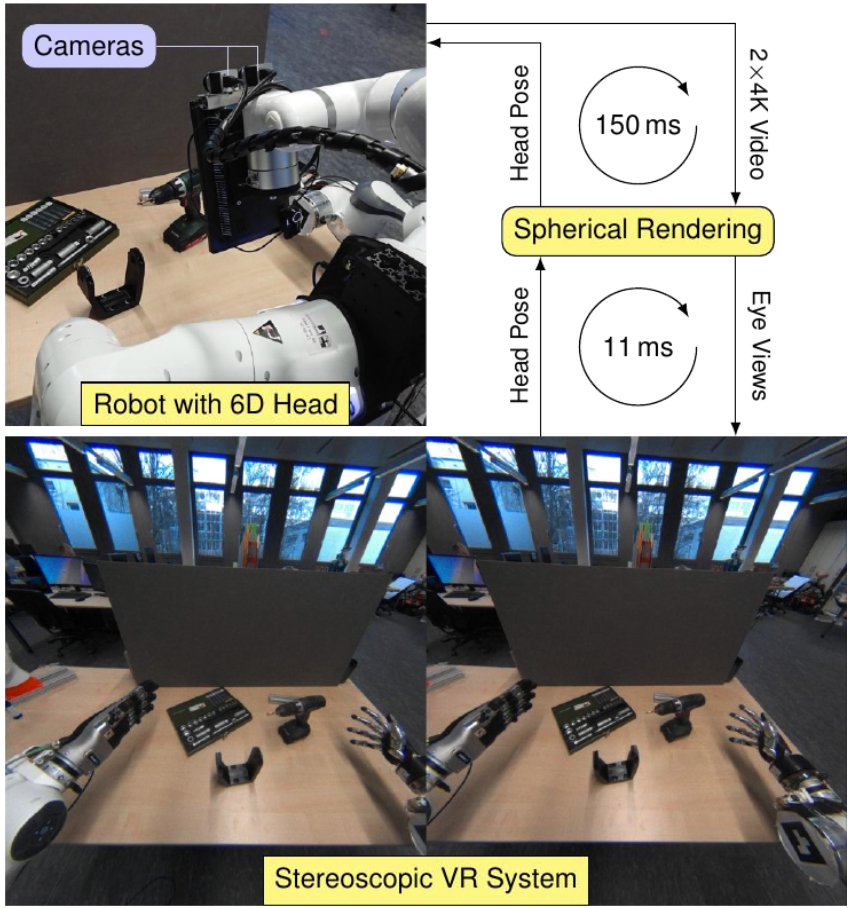
Tasks

1. Make a coffee
2. Greet the recipient
3. Measure temperature
4. Measure blood pressure
5. Measure oxygen saturation
6. Help recipient with jacket

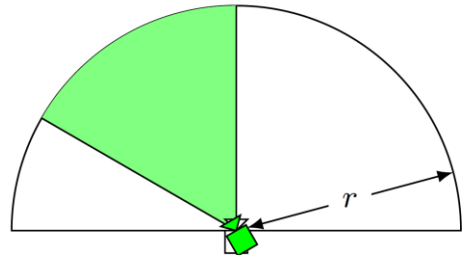


[Schwarz et al. IROS 2021]

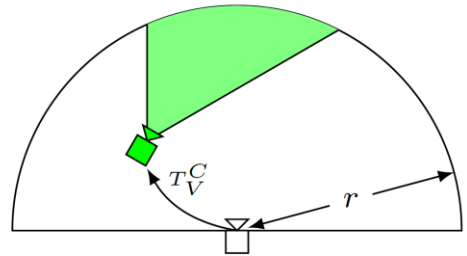
NimbRo Avatar: Immersive Visualization



- 4K wide-angle stereo video stream
- 6D neck allows full head movement
  - Very immersive
  - Good hand-eye coordination
- Spherical rendering technique hides movement latencies
  - Assumes constant depth



Exact for pure rotations

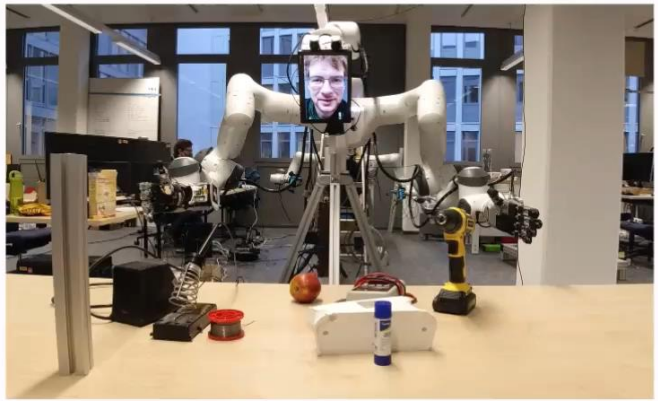


Distortions for translations

[Schwarz and Behnke Humanoids 2021]

NimbRo Avatar: Immersive Visualization

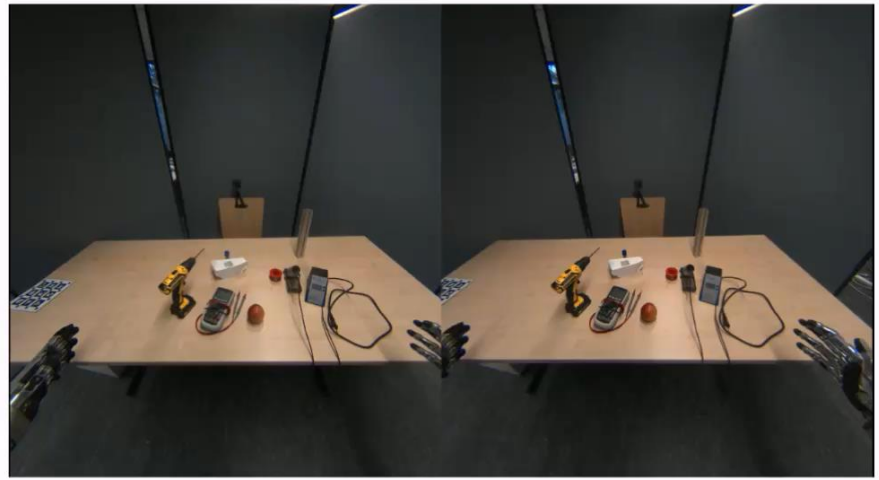
Avatar Robot



Wide-Angle Stereo

Operator

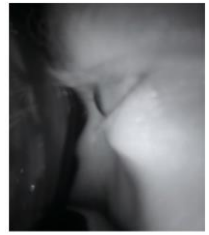
HMD View



[Schwarz and Behnke Humanoids 2021]

NimbRo Avatar: Operator Face Animation

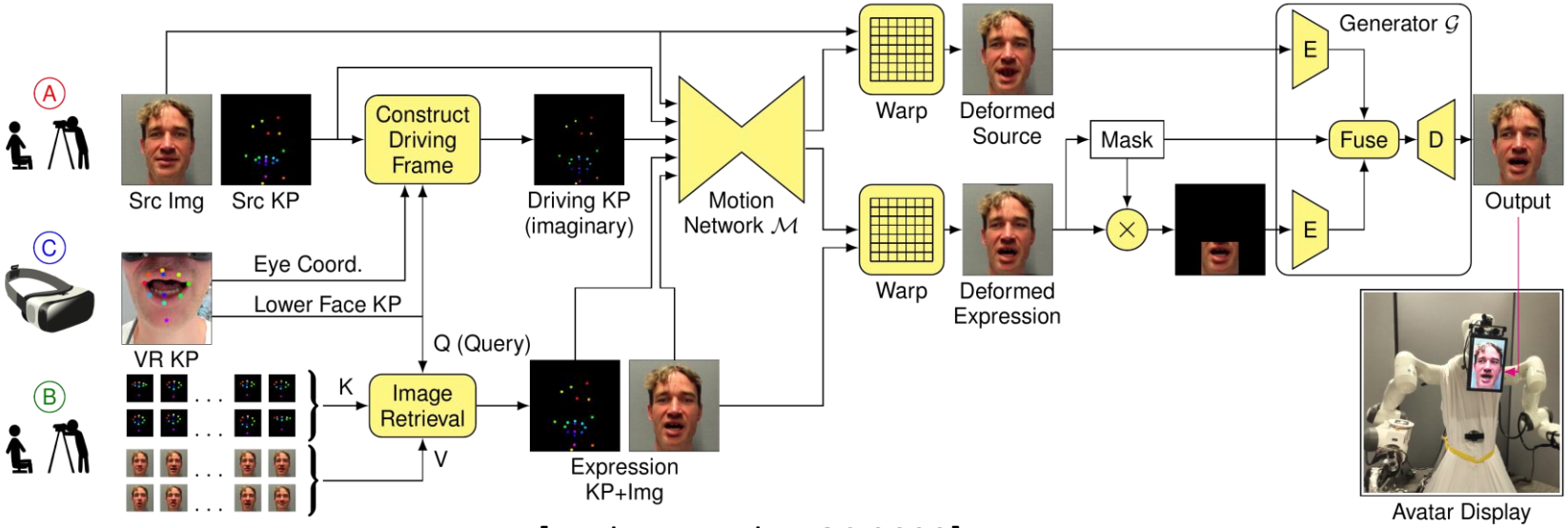
- Operator images without HMD
- Capture mouth and eyes
- Estimate gaze direction and facial keypoints


Left Eye

Mouth

Right Eye

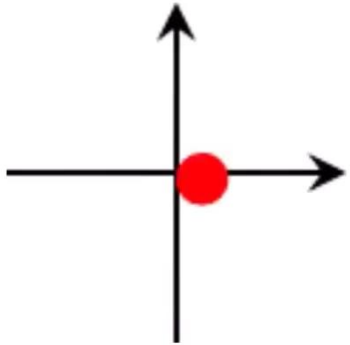
- Generate animated operator face using a warping neural network



[Rochow et al. IROS 2022]

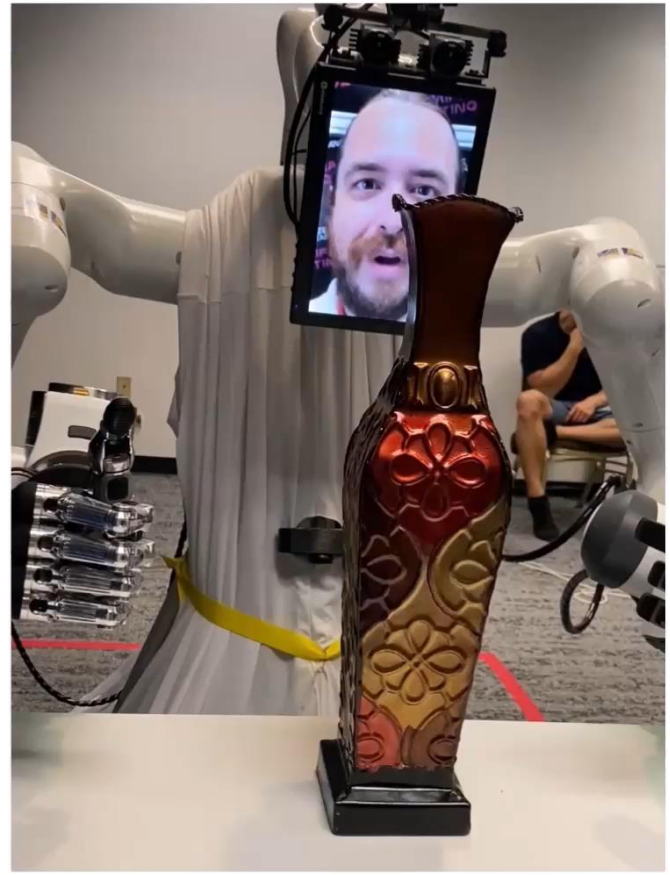
NimbRo Avatar: Operator Face Animation

Gaze Direction



Mouth Cam

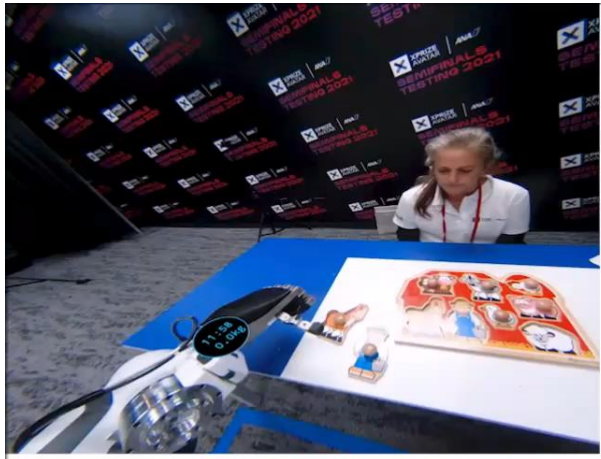
Output

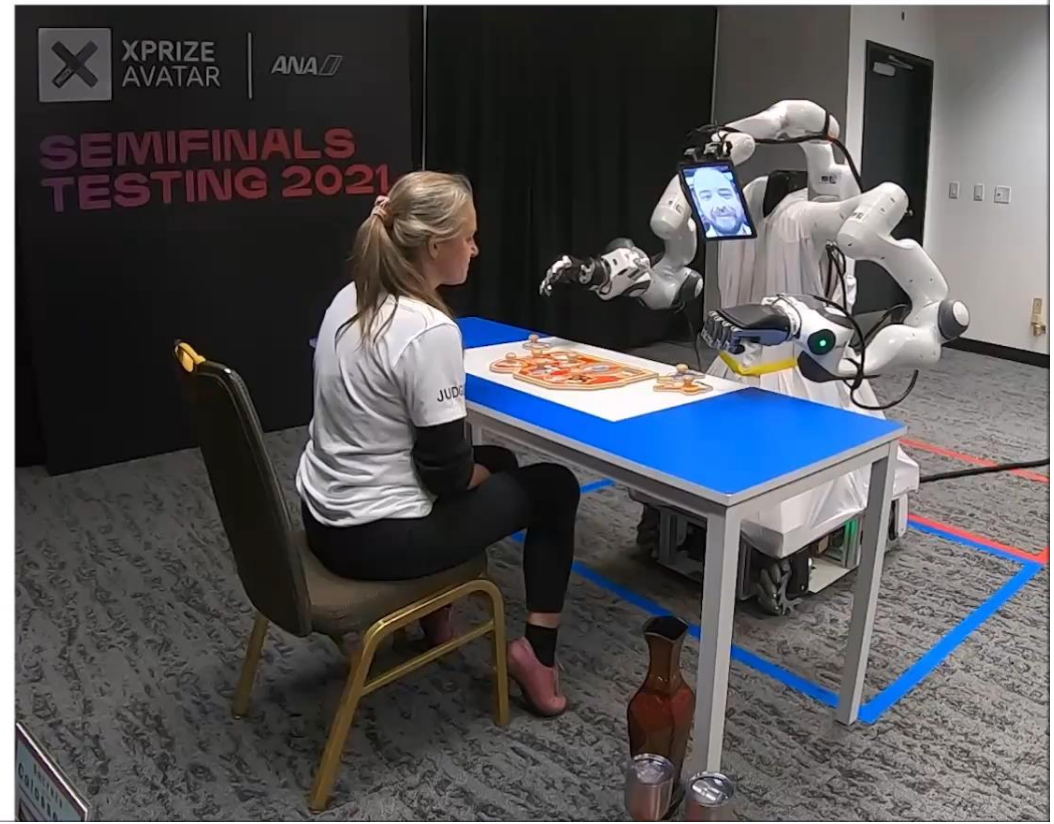


[Rochow et al. IROS 2022]

NimbRo Avatar

Avatar XPRIZE Semifinals





[Schwarz et al. IROS 2021]

Semifinals Conclusions

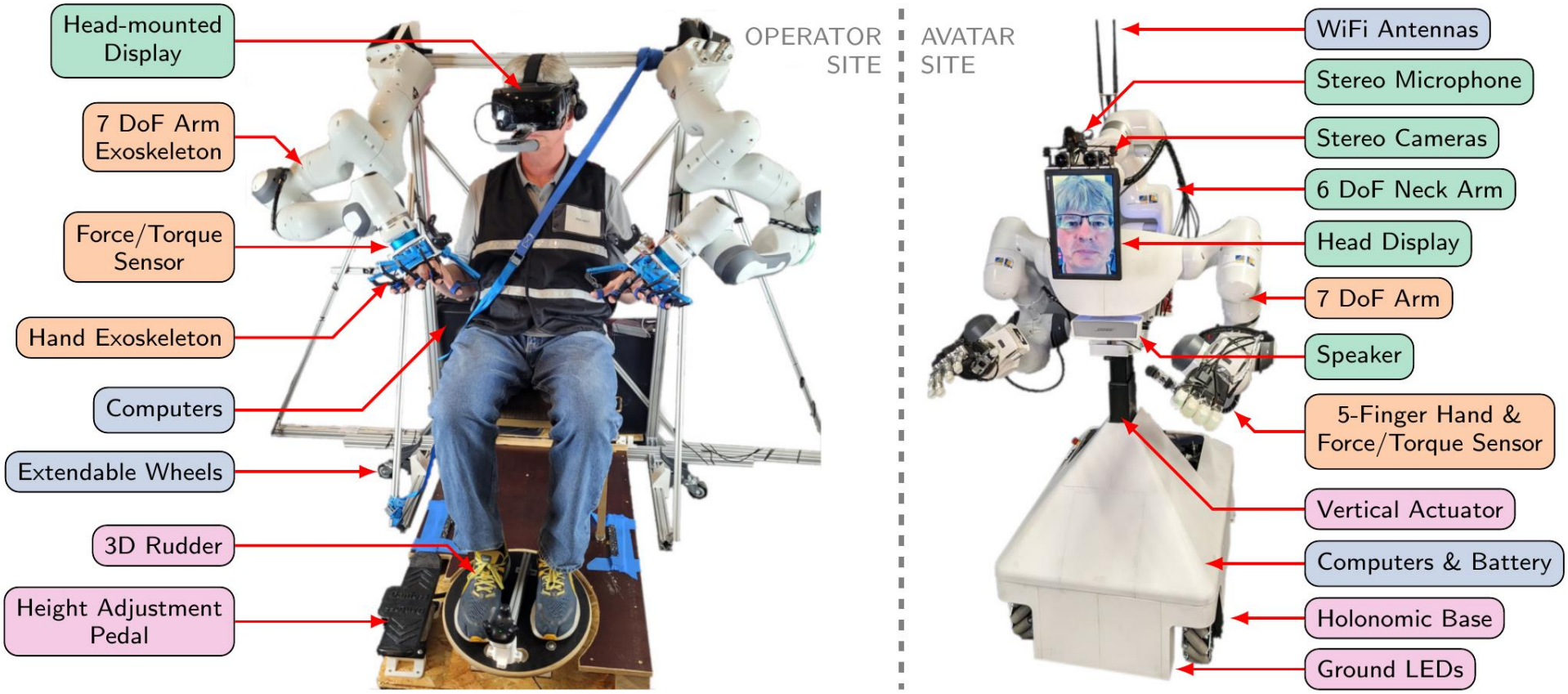
- Designed an Avatar system for intuitive immersive telepresence
- Very good immersive visualization
- Operator-Recipient interaction with facial animation
- Bimanual human-like manipulation with force and haptic feedback
- Omnidirectional drive with birds-eye navigation view
- Scored 99/100 points, ranked 1st in the Semifinals
- Judges seemed to enjoy our system

New Finals Requirements

ANA AVATAR XPRIZE

- Untethered avatar robot, more mobility
- Movable operator station
- 10 tasks in a sequence, including haptics
- System reliability extremely important
- Tasks fulfillment has highest importance in scoring
- Subjective criteria also important
- Trial time to break ties

NimbRo Avatar Finals System

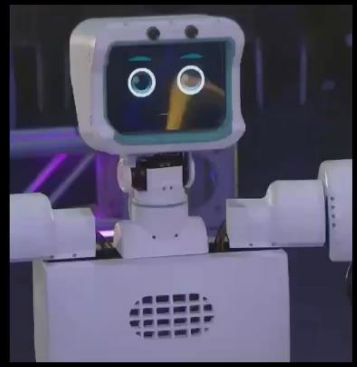


[Lenz et al. 2023, submitted to International Journal of Social Robotics]

Finals Test Run Day 1

Face Animation @ Finals

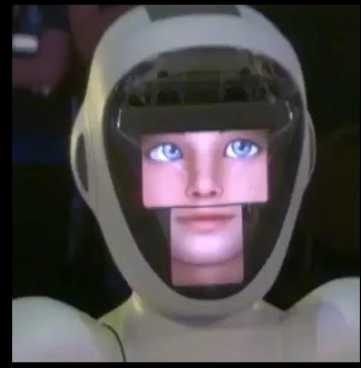
Team UNIST



Ours (NimbRo)

Team AVATRINA [13]

Northeastern [12]



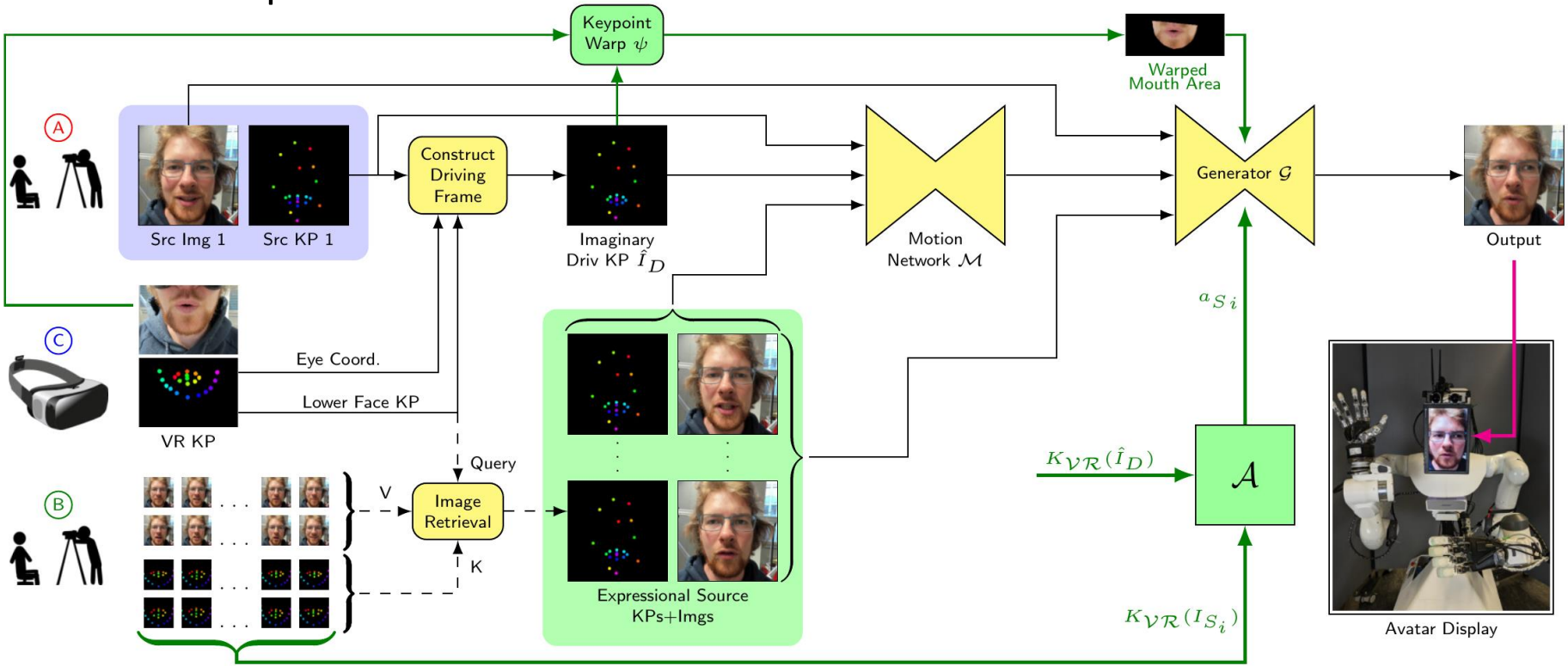
i-BOTICS

Pollen Robotics

Source: Official XPRIZE Avatar live stream

Improved Operator Face Animation

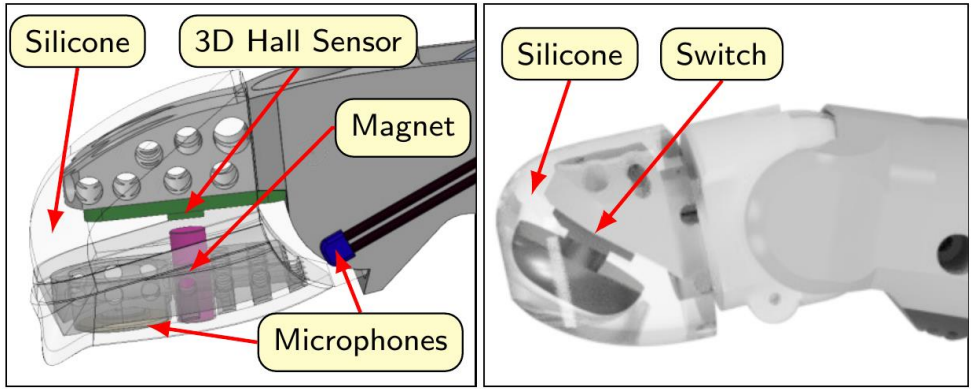
- Better temporal continuity
- Direct incorporation of mouth video



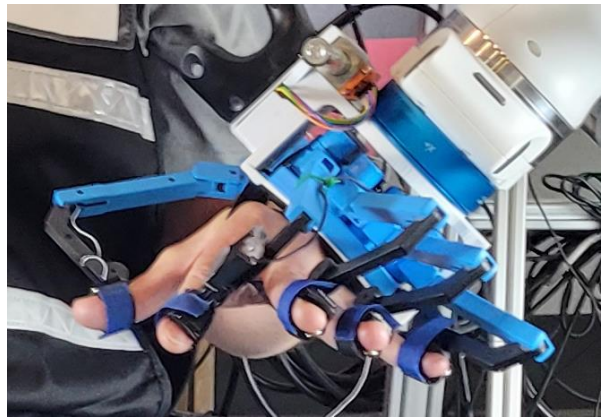
[Rochow et al. 2023, submitted to IROS]

Haptic Perception

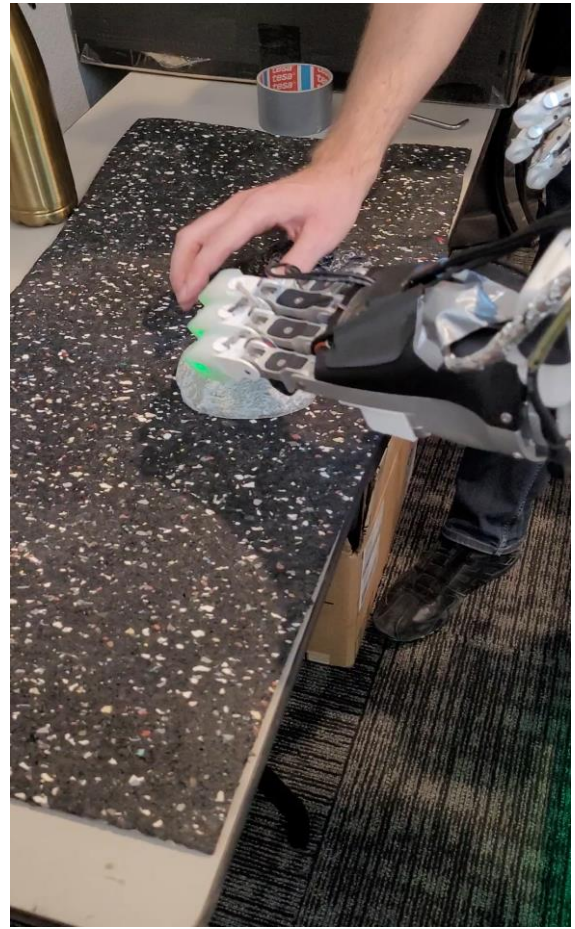
- Sensors in the finger tips



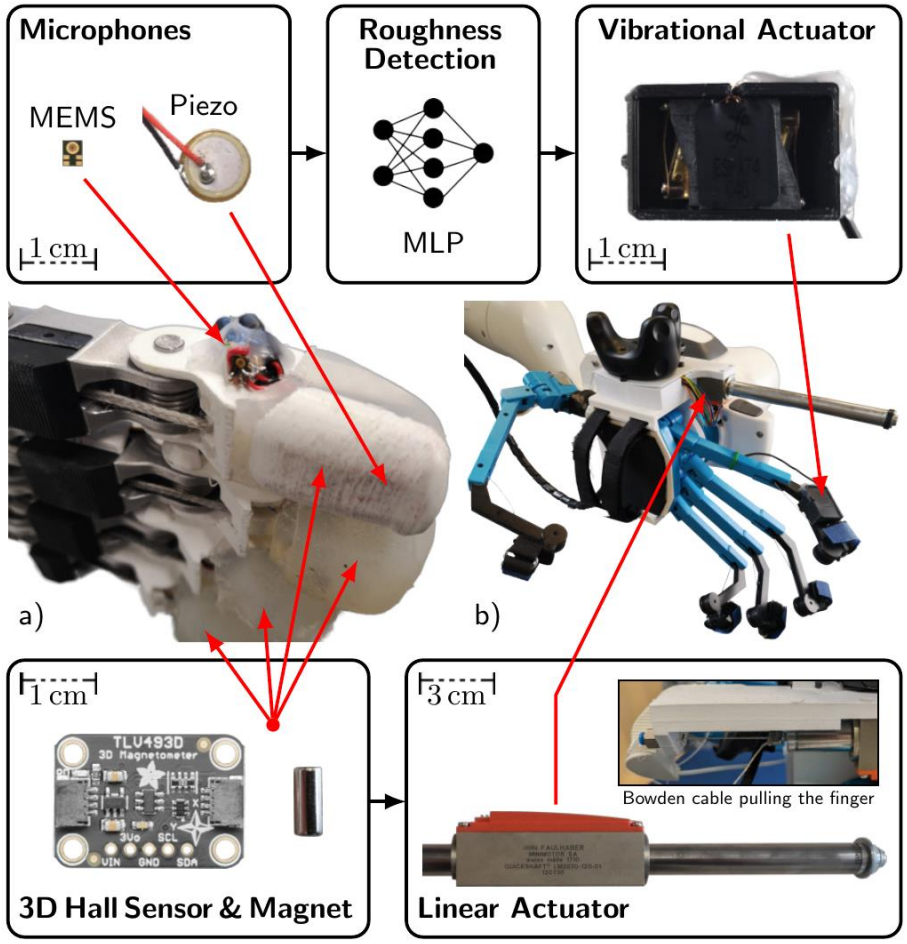
- Actuators on the hand exoskeleton



[Pätzold et al. SMC 2023]



Roughness Perception



Dataset of
rough and smooth objects

[Pätzold et al. SMC 2023]

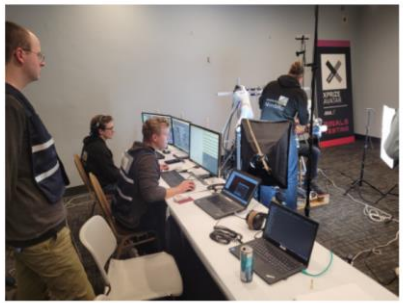
[Lenz et al. 2023, submitted to International Journal of Social Robotics]

Operator Training

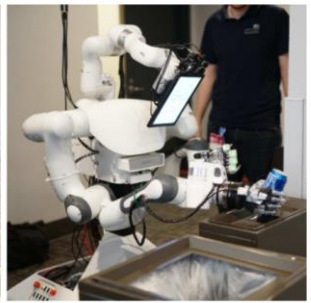
(a) Introduction

(b) Locomotion

(c) Grasping



(d) Monitoring crew



(e) Free experiments

Training	Time [min]
System overview	3
Face animation video w/o HMD	2
Put on HMD	1
Face animation video with HMD	2
Strap in hands	4
Enable arm and hand control	3
Locomotion training (T1, T5, T8)	4
Training switch and canister (T4, T6, T7)	5
Training power drill (T9)	5
Training stones (T10)	10
Enjoy the system	3
System recovery & recap	3
Total training	45

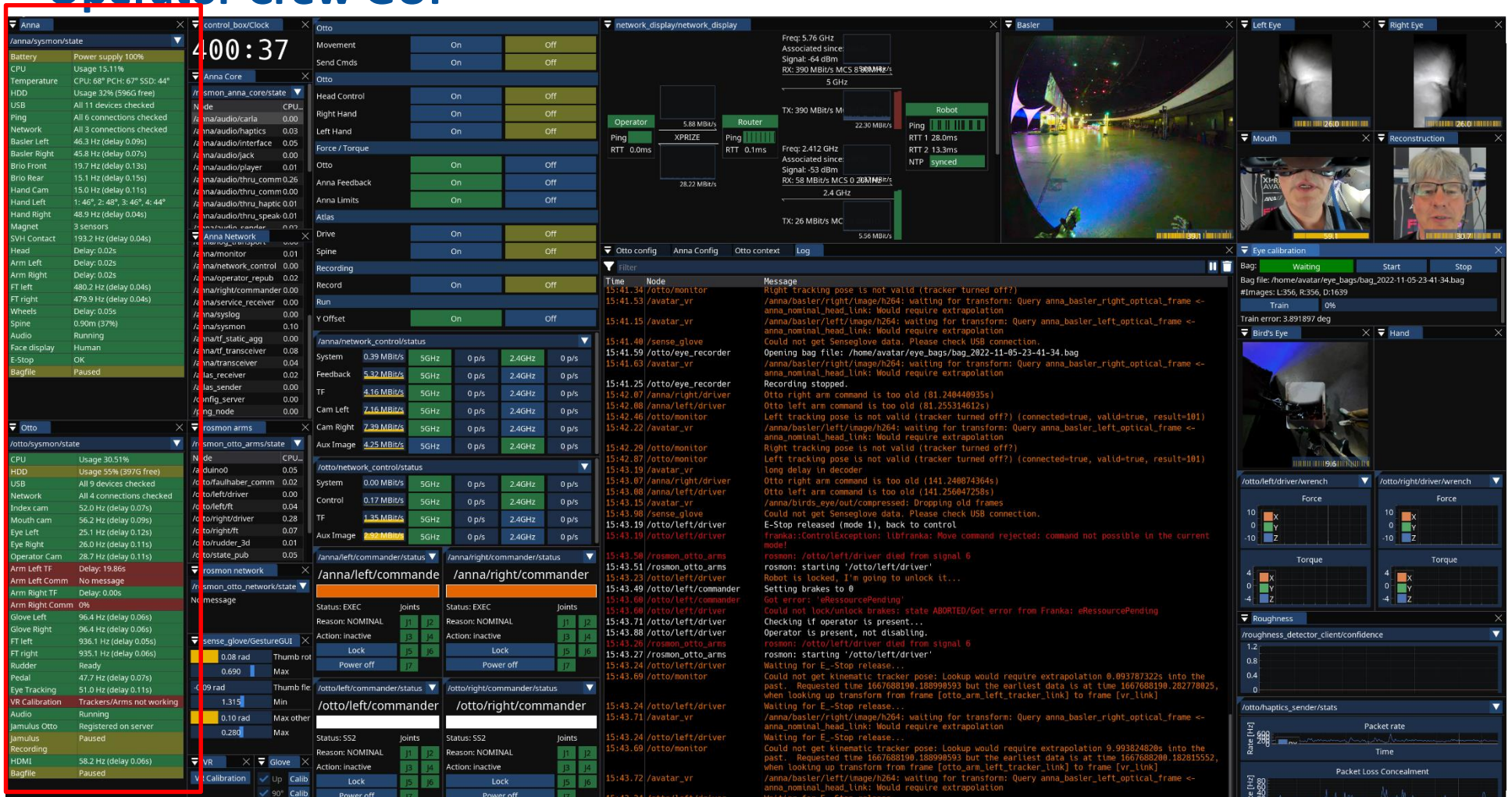
- Dedicated roles: Communication with operator, Software control, Face animation, Hardware support
- Trade-off learning by doing vs. explicit instruction

[Lenz et al. 2023, submitted to International Journal of Social Robotics]

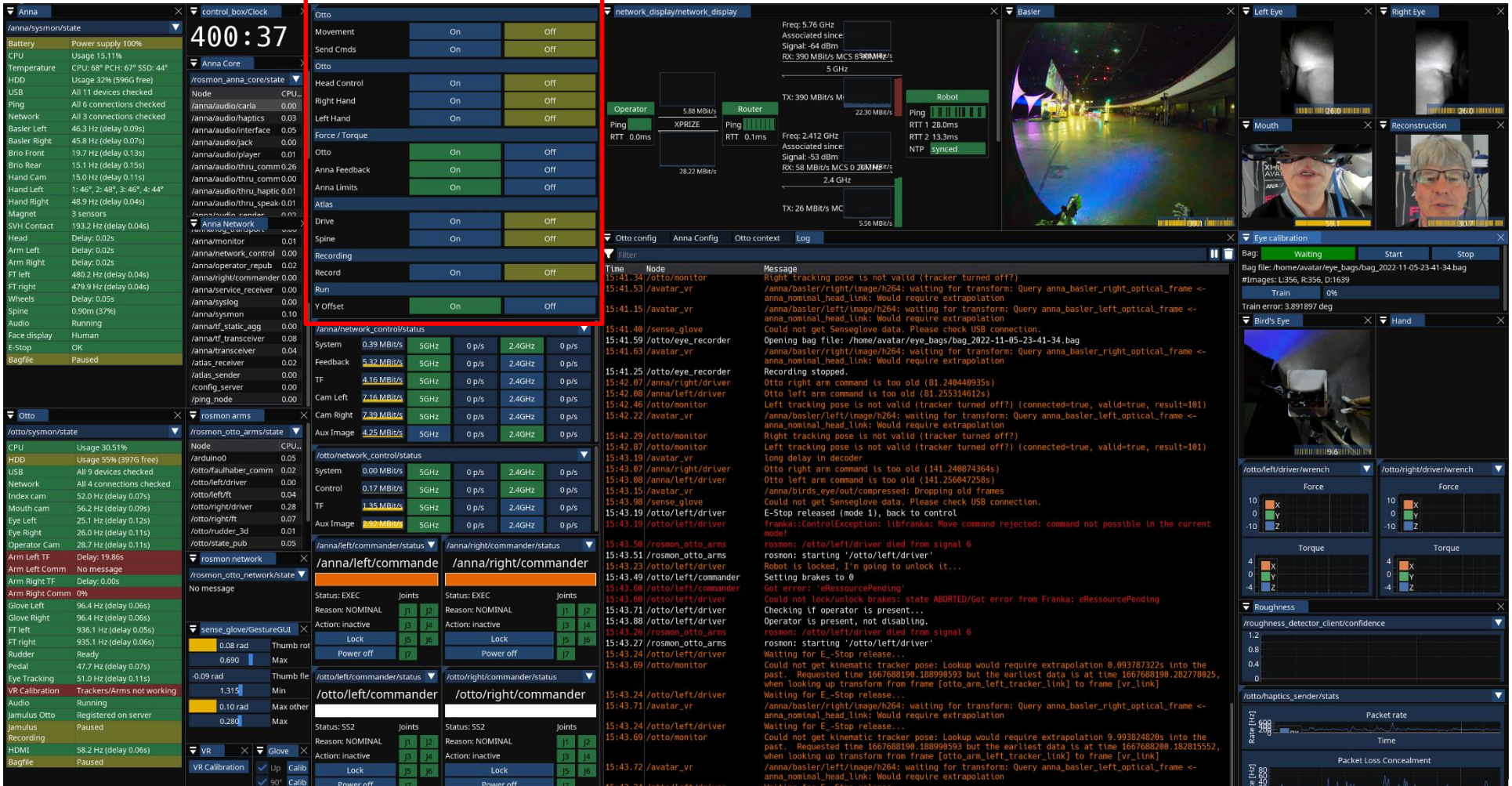
Moving into the Arena

- Seamless roaming / disconnection handling with UDP data streaming
- No calibration/initialization/button press
- Essential: Operator room crew in the voice loop during setup
- Gamepad control
- To quote Colin: Hydrate to Dominate!

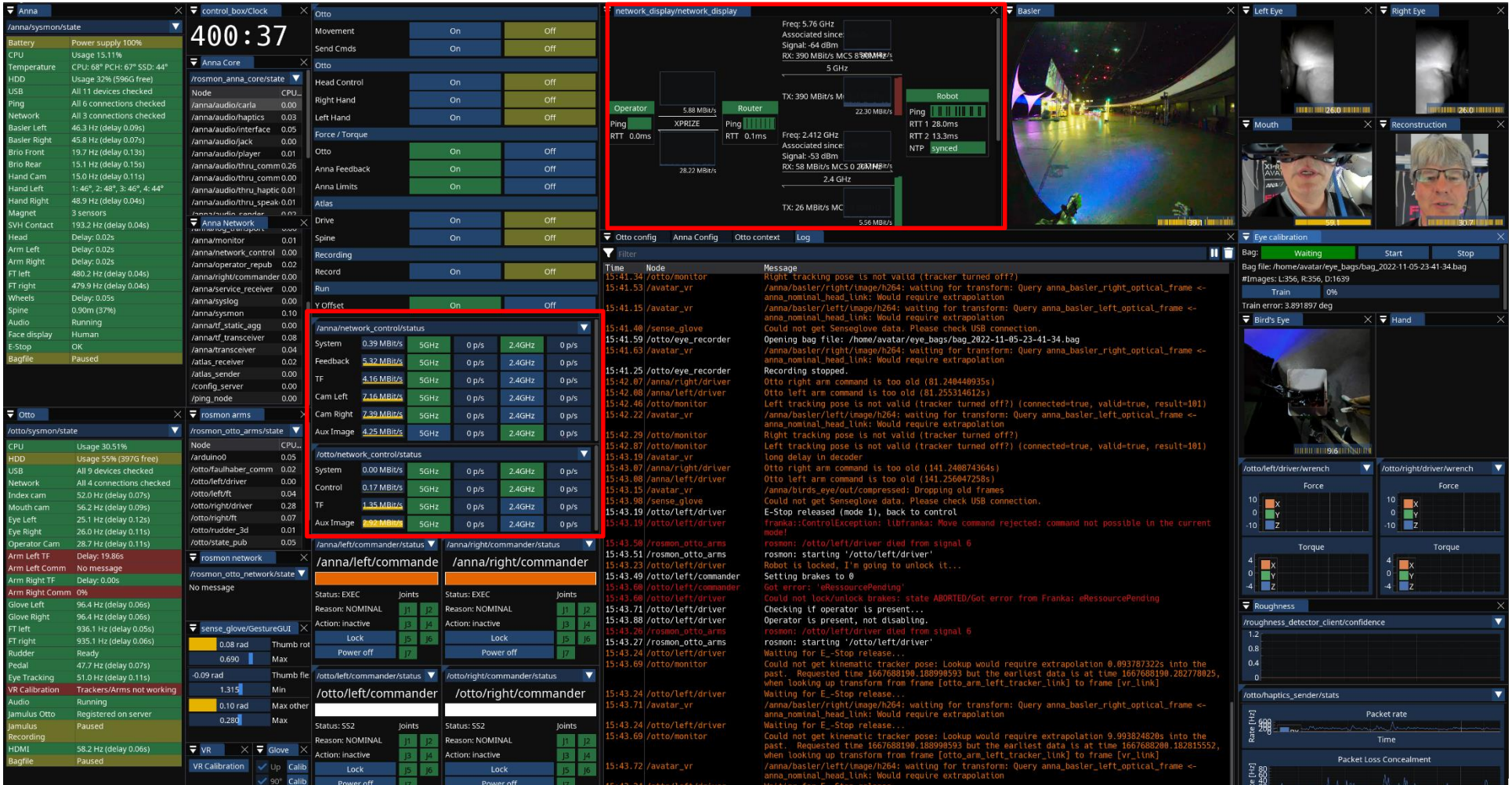
Operator Crew GUI



Operator Crew GUI

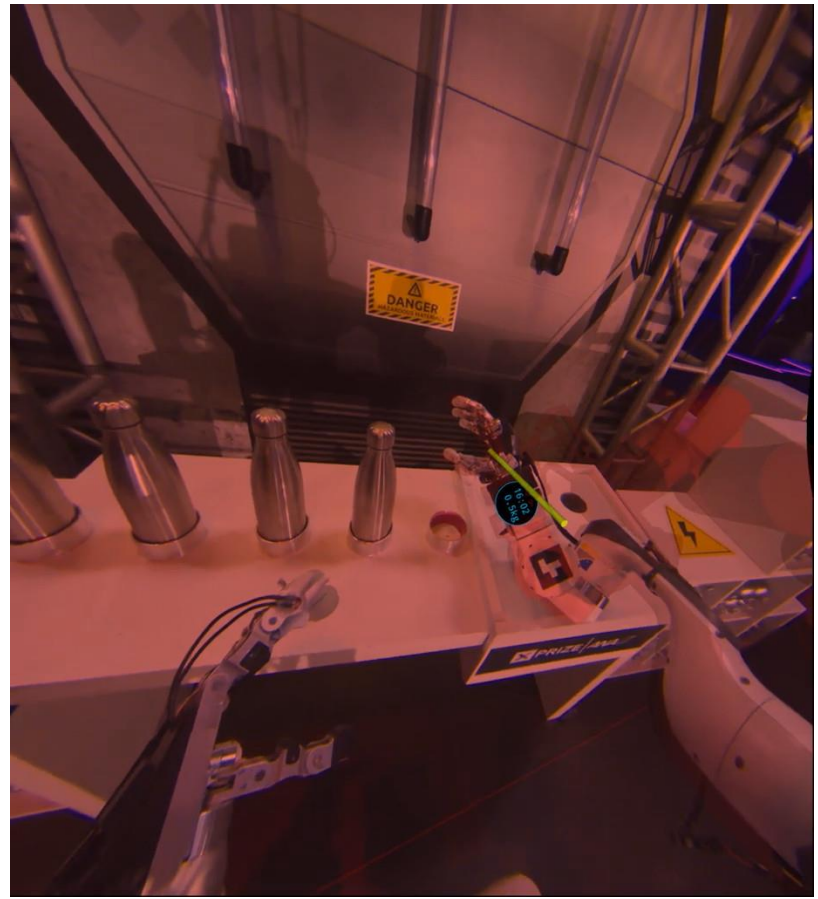


Operator Crew GUI



Reliability Features

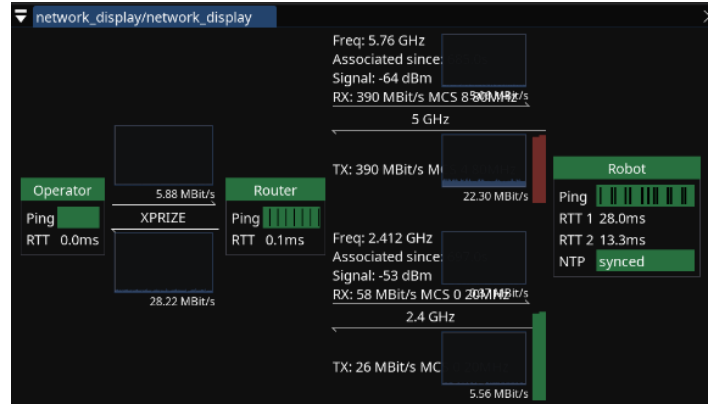
1. Operator crew awareness
2. Automatic arm resets
3. ROS node respawn
4. State- and connectionless network system (pure UDP)
5. Redundant WiFi connections
6. PC Watchdog



[Schwarz et al. 2023, arXiv:2303.03297, submitted to IROS]

Network Details

- Separate ROS cores operator / avatar
- Pure UDP, no re-connect / initialization
- Main camera stream (stereo 2472×2178 @46 fps) is HEVC-encoded & decoded on GPU (NVENC).
  Total Bandwidth: ~14 MBit/s
- Control data is sent redundantly
- Monitor packet loss due to congestion
- The core software is already open source, more to come:
  https://github.com/AIS-Bonn/nimbro_network



Audio Details

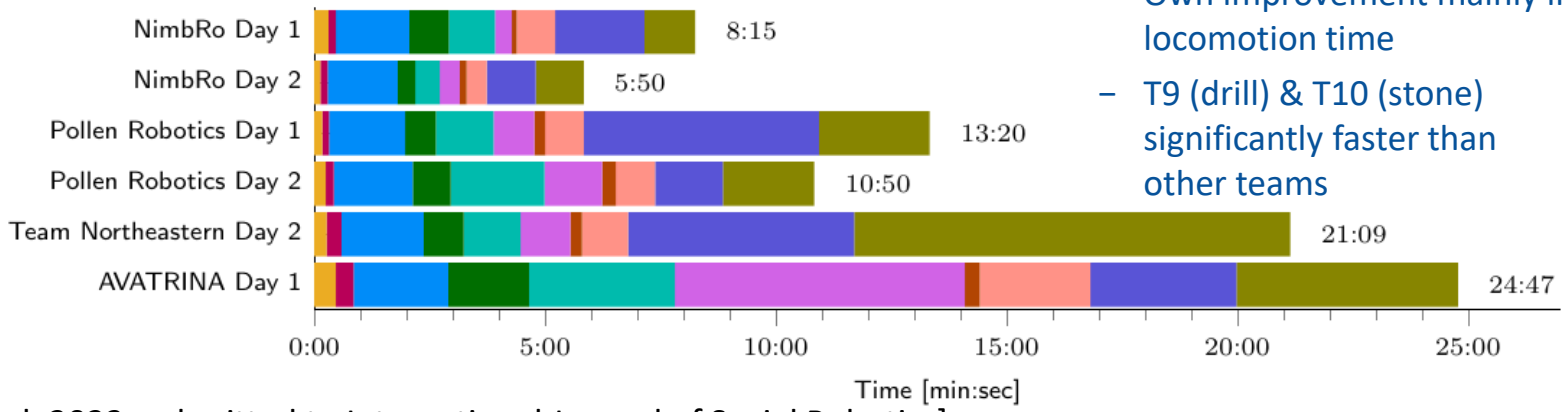
- Low latency solution utilizing the *JACK Audio Connection Kit*
- Redundant UDP transmission via the *OPUS audio codec*
- *NVIDIA MAXINE* for GPU-accelerated *acoustic echo cancelation*
- *jamulus* for team communication with operator and recipients

[Schwarz et al. 2023, arXiv:2303.03297, submitted to IROS]

Finals Day 2 Testing

Finals Timings

Team	Day	Time [mm:ss] Start	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	Total
NimbRo	1	00:00	00:18	00:10	01:35	00:52	01:00	**00:22**	**00:06**	00:50	01:56	01:06	08:15
	2	00:00	**00:08**	**00:09**	01:31	**00:23**	**00:32**	00:26	00:09	**00:26**	**01:04**	**01:02**	**05:50**
	1→2	0:00	-0:10	-0:01	-0:04	-0:29	-0:28	+0:04	+0:03	-0:24	-0:52	-0:04	-2:25
Pollen Robotics	1	00:00	00:10	**00:09**	01:39	00:40	01:15	00:53	00:14	00:50	05:06	02:24	13:20
	2	00:00	00:15	**00:09**	01:43	00:49	02:02	01:15	00:18	00:51	01:28	01:59	10:50
Team Northeastern	1	00:00	00:33	00:24	02:08	01:43	04:03	01:27	00:36	01:56			12:50
	2	00:00	00:16	00:19	01:47	00:52	01:14	01:05	00:15	01:00	04:54	09:27	21:09
AVATRINA	1	00:00	00:28	00:23	02:03	01:45	03:10	06:17	00:19	02:24	03:10	04:48	24:47
	2	00:00	00:24	00:12	01:39	01:05	02:50	00:48	00:11	01:30	02:43		11:22
i-Botics	1	00:00	00:13	00:26	**01:23**	01:53	01:57	01:52	02:07	02:57	09:47		22:35
	2	00:00	00:19	00:12	01:36	03:25							05:32



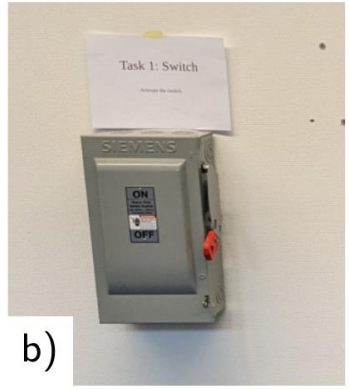
- Own improvement mainly in locomotion time
- T9 (drill) & T10 (stone) significantly faster than other teams

[Lenz et al. 2023, submitted to International Journal of Social Robotics]

Team NimbRo

User Study

a)



b)

c)

d)

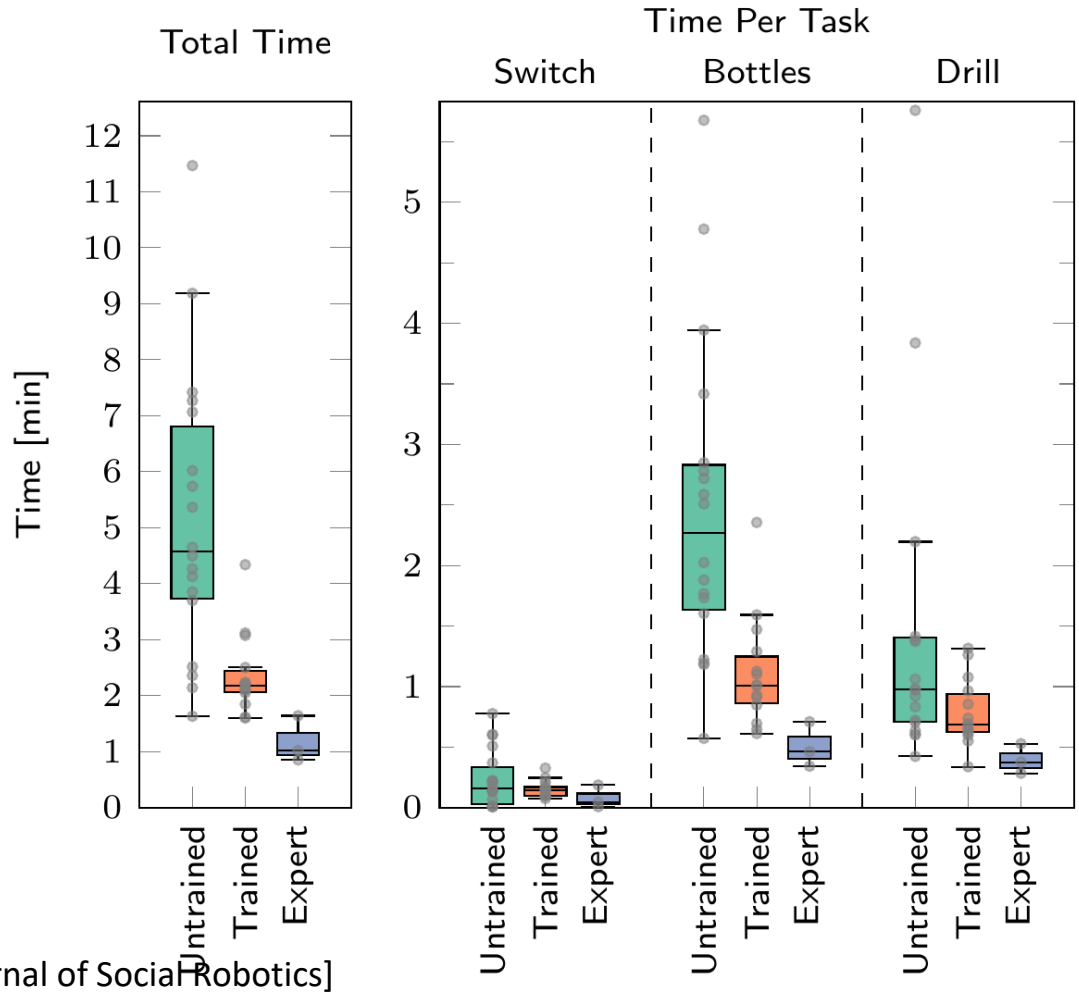
- Three tasks, similar to finals (Locomotion+Switch, Bottles, Drill)
- 35 participants, 32 with no prior experience of the system
- All participants: 2 min intro video explaining the system (task agnostic)
- Three groups: No further training, 10 min task training, expert team members
  similar to Finals

[Lenz et al. 2023, submitted to International Journal of Social Robotics]

User Study

- Unsurprising: Clear advantage of training (2x)
- Unsurprising: Expert operators are very fast (2x)
- Untrained operators could still solve all tasks in reasonable time
- All participants were able to solve the tasks

=> System is very intuitive, but short instruction on tasks improves completion time.



[Lenz et al. 2023, submitted to International Journal of Social Robotics]

Lessons Learned

- Robustness is key
- Latency is the enemy
- Frequent testing under competition conditions: system & people!
- 1:1 correspondence is best
- 6D head motion actually simplifies manipulation control
- Immersive control overlays – don’t break immersion!
- Facial animation and gestures: Head & gaze direction enables shared awareness
- Modified components

[Lenz et al. 2023, submitted to International Journal of Social Robotics]

Questions?

Join the Winning Team!

We are hiring.