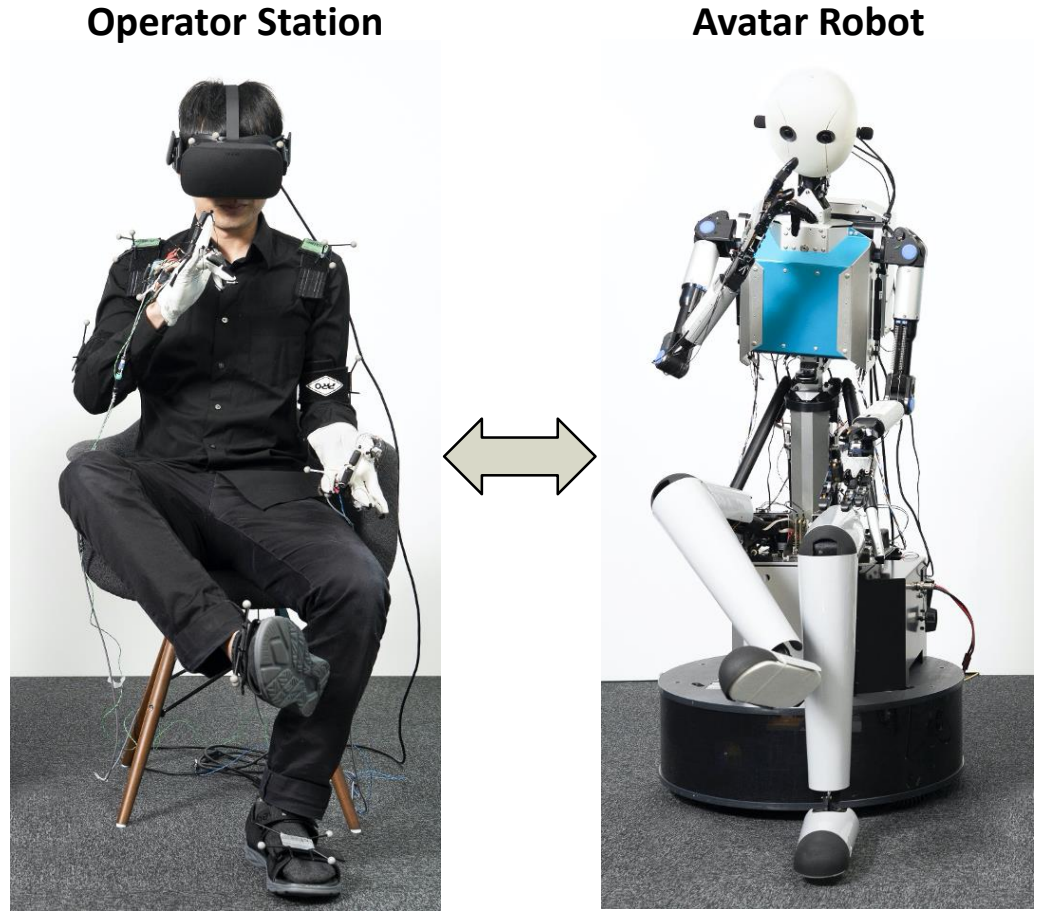
The NimbRo Immersive Telepresence System Winning the ANA Avatar XPRIZE Finals: Human-Centric Evaluation and Lessons Learned

Sven Behnke

University of Bonn
Computer Science Institute VI –
Intelligent Systems and Robotics
Autonomous Intelligent Systems

Telepresence Systems

- Enable a human operator to be present at a remote location
- Capture remote location with cameras, microphones, force & haptic sensors, etc.
- Display remote measurements to the operator
- Capture operator movements, speech, and expressions
- Transfer them to avatar robot



[TELESAR VI, Tachi et al. IJHR 2020]

Telepresence Applications

- Remote visits to family and friends
- Business trips
- Health care
- Personal assistance
- Remote work
- Disaster response
- Space
- Underwater
- Remote driving
- Many more …

[Hung et al. 2023]

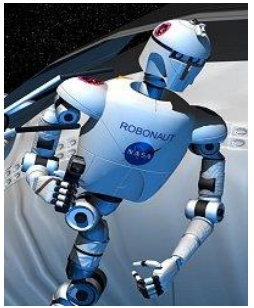
[OhmniLabs Ohmni]

[Intuitive Da Vinci]

[Pollen Reachy]

[Telexistence]

[KAIST DRC Hubo]



[NASA Robonaut]

[Stanford OceanOneK]

[Fetch]

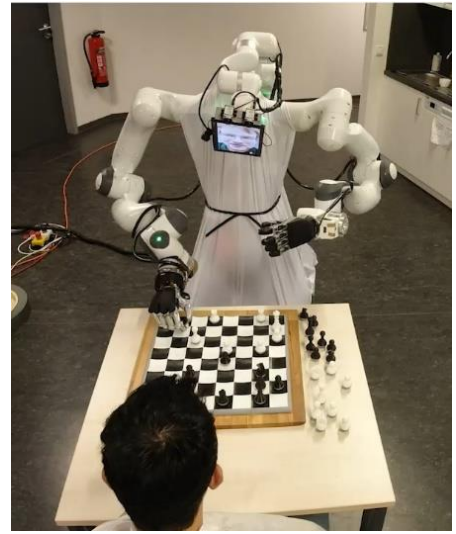
Experience with Teleoperated Robots

- Multiple domains
- Often motivated by competitions and challenges

RoboCup@Home

DARPA Robotics Challenge
DLR SpaceBot Cup

CENTAURO

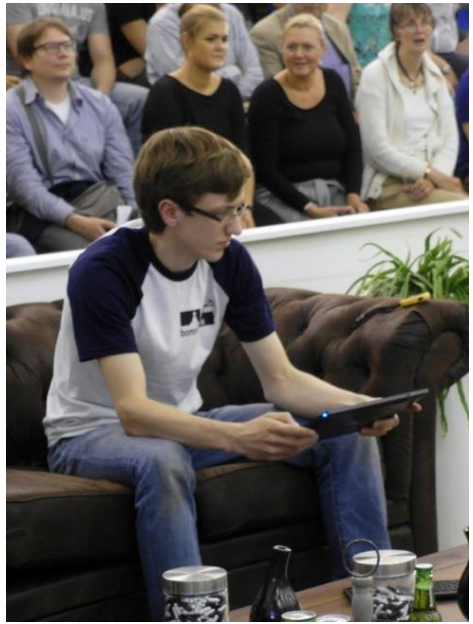


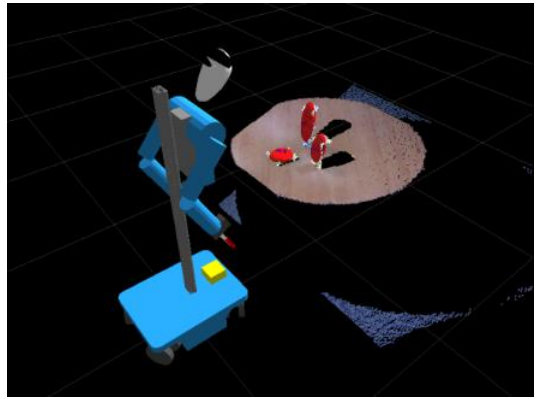
ANA Avatar XPRIZE

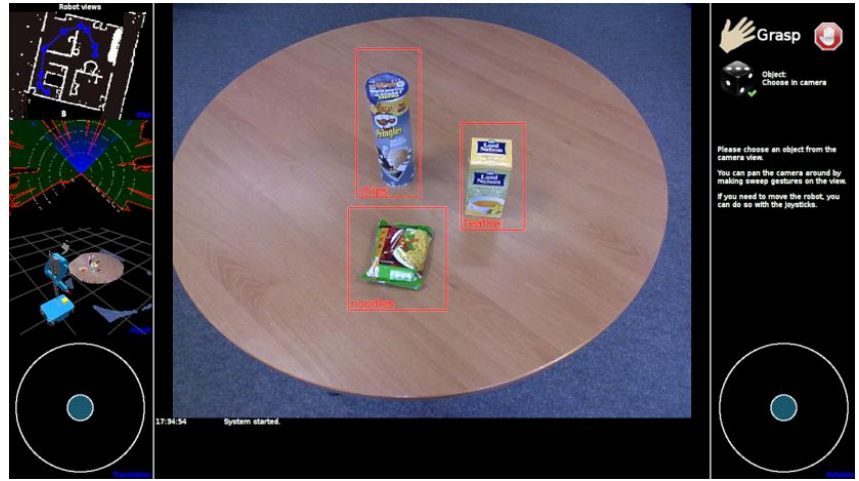
Cognitive Service Robot Cosero

Handheld Teleoperation Interface

- Three levels of autonomy/control:
  - Task level: Get me a beer!
  - Skill level: Grasp, place, navigate, …
  - Direct control: Locomotion, manip.





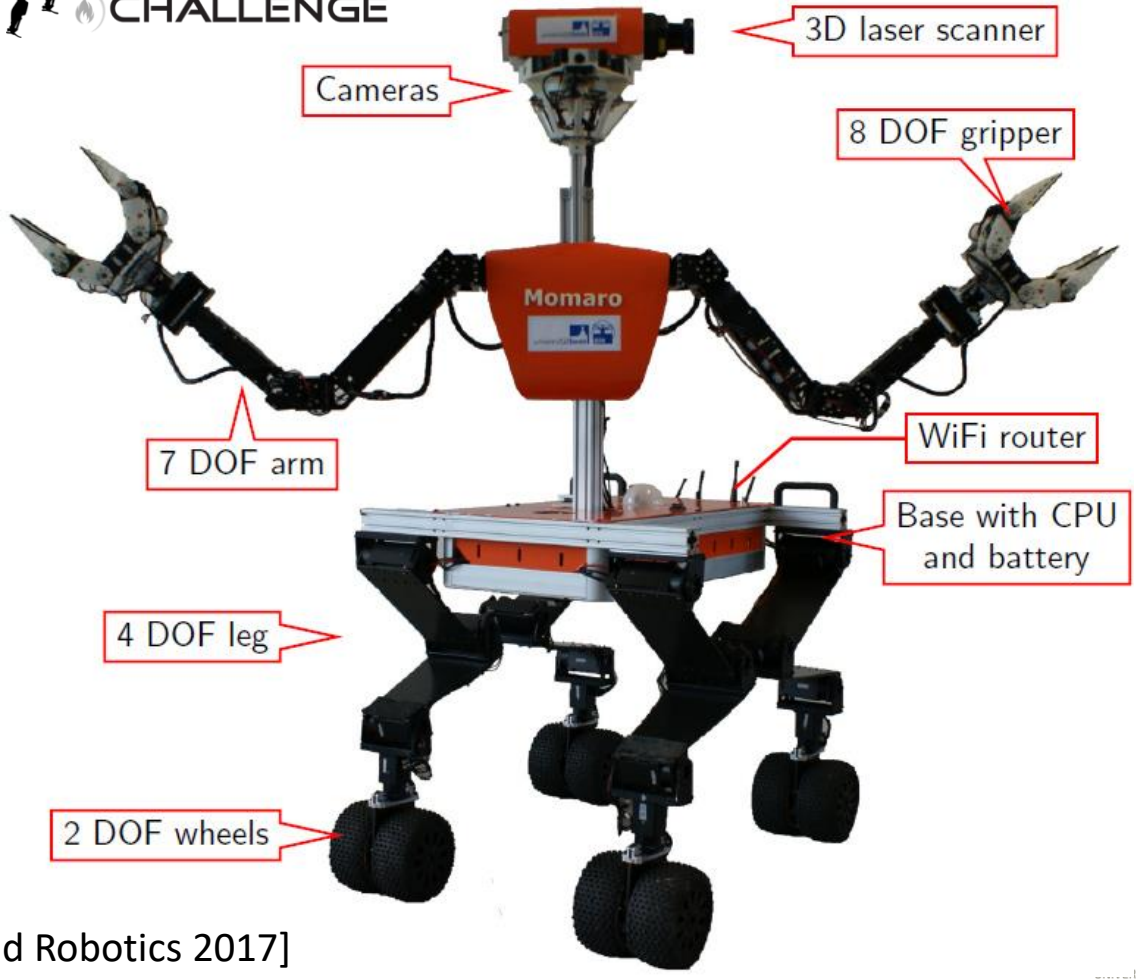


[Schwarz, Stückler, Behnke, HRI 2014]

Mobile Manipulation Robot Momaro

DARPA ROBOTICS CHALLENGE

- Four compliant legs ending in pairs of steerable wheels
- Anthropomorphic upper body
- Sensor head
  - 3D laser scanner
  - IMU, cameras

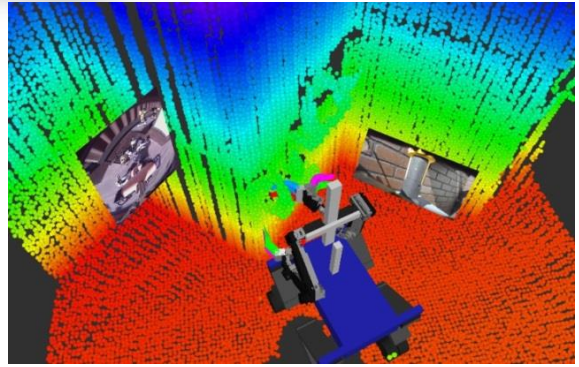


[Schwarz et al. Journal of Field Robotics 2017]

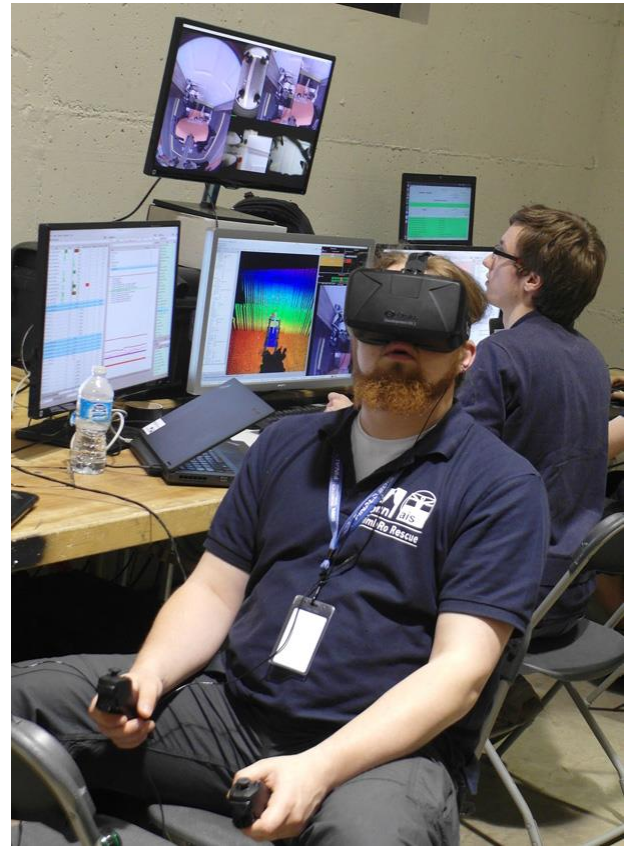
Manipulation Operator Interface

- 3D head-mounted display

- 3D environment model + images



- 6D magnetic tracker



[Rodehutskors et al., Humanoids 2015]

DARPA Robotics Challenge

[Schwarz et al. Journal of Field Robotics 2017]

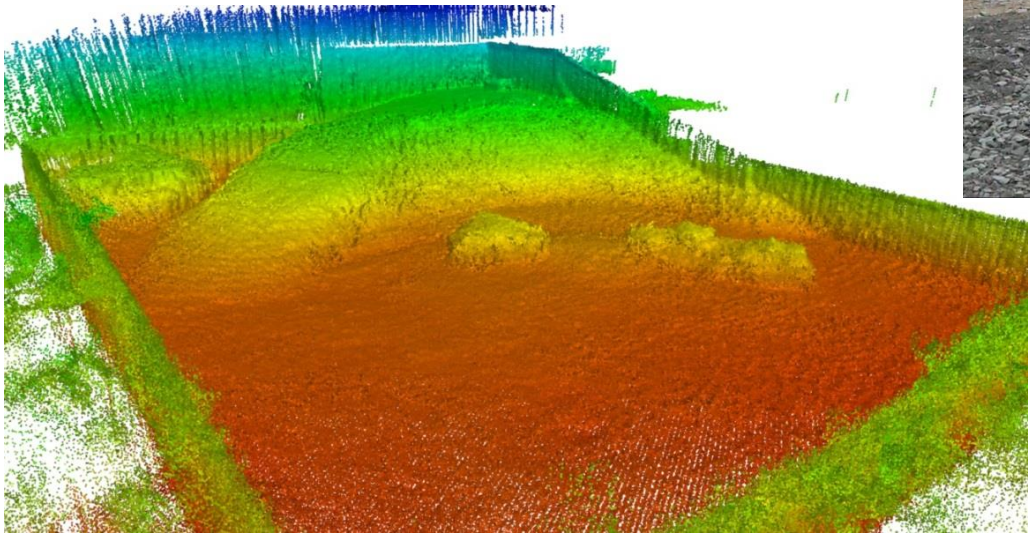
Team NimbRo Rescue

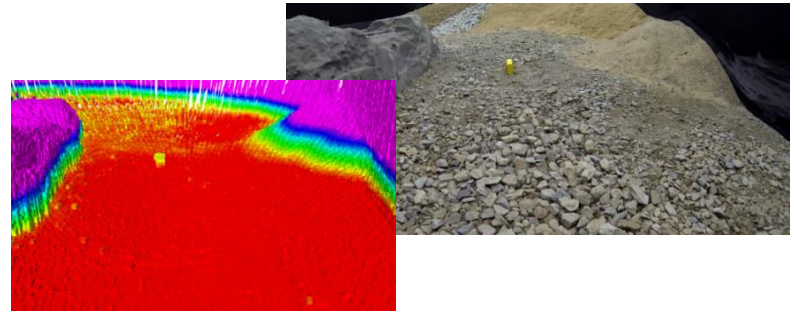
Best European Team (4th place overall), solved seven of eight tasks in 34 minutes

[Schwarz et al. Journal of Field Robotics 2017]

DLR SpaceBot Cup 2015

- Mobile manipulation in rough terrain



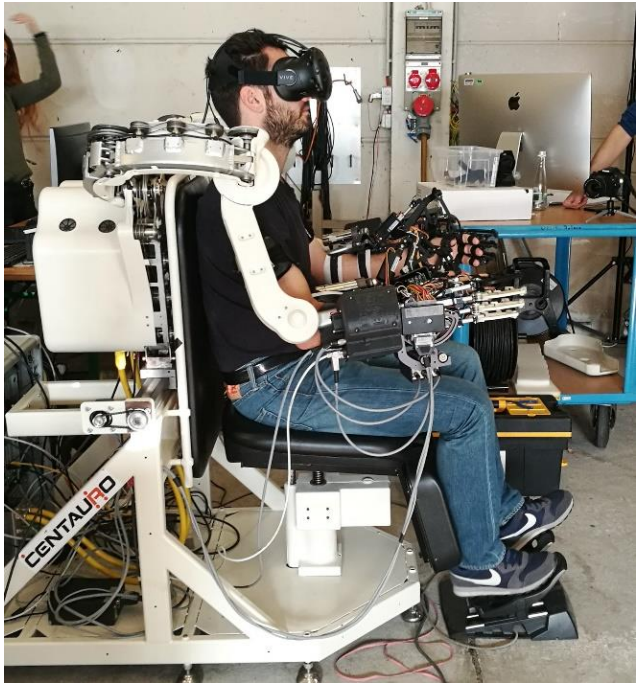


[Schwarz et al., Frontiers on Robotics and AI 2016]

[Schwarz et al., Frontiers on Robotics and AI 2016]

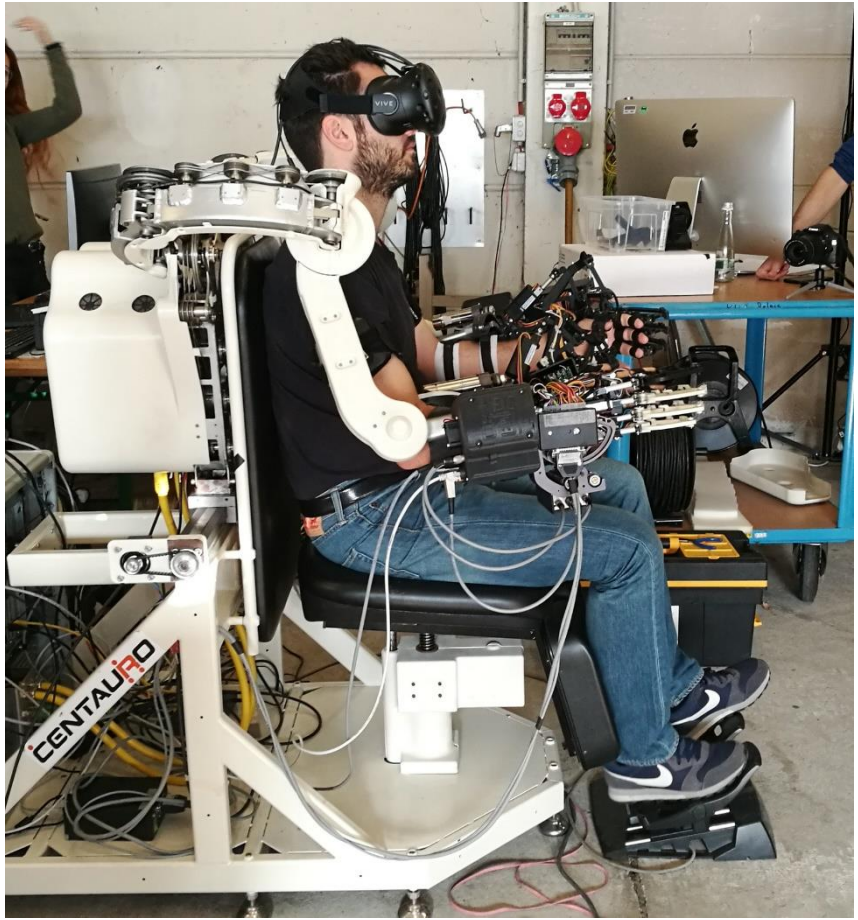
Robust Mobility and Dexterous Manipulation in Disaster Response by Fullbody Telepresence in a Centaur-like Robot

- Four-legged robot with steerable wheels and anthropomorphic upper body
- Immersive teleoperation through exoskeleton with HMD



[Klamt et al., Journal of Field Robotics 2020]

Immersive Operator Interface



Stereo Visual and Audio Feedback
- Head-Mounted Display
- Audio Headset

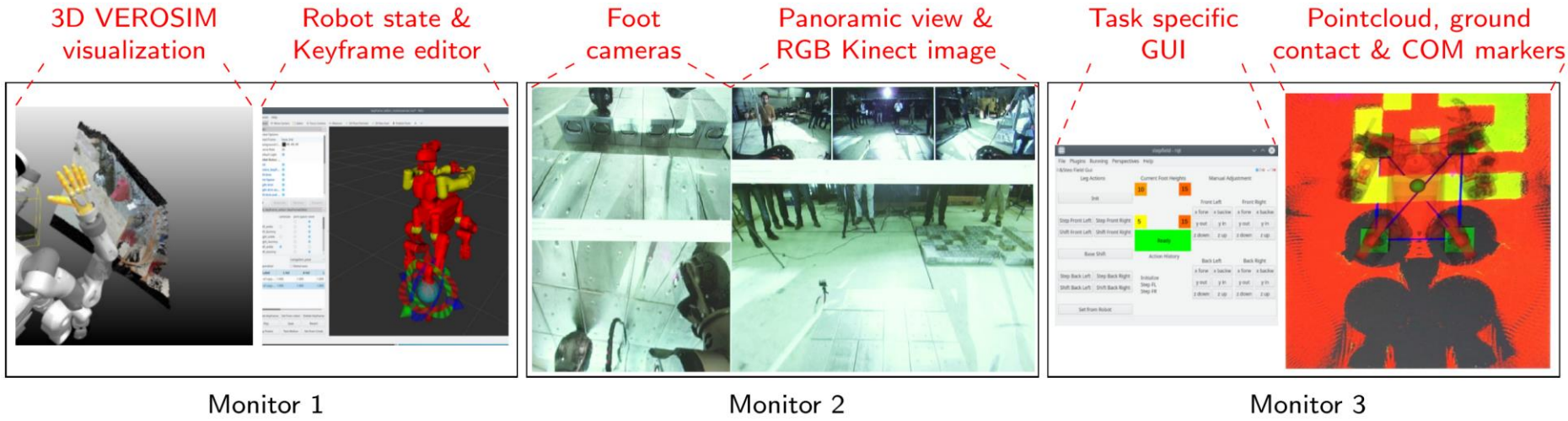
Arm-Hand Movements
- Arm and wrist exoskeleton
- 7 DoFs, wide workspace
- Force and torque feedback

Grasping Control
- Hand exoskeleton
- Independent Fingers
- Force feedback

Navigation
- 3 DoFs Pedals

[Klamt et al., Journal of Field Robotics 2020]

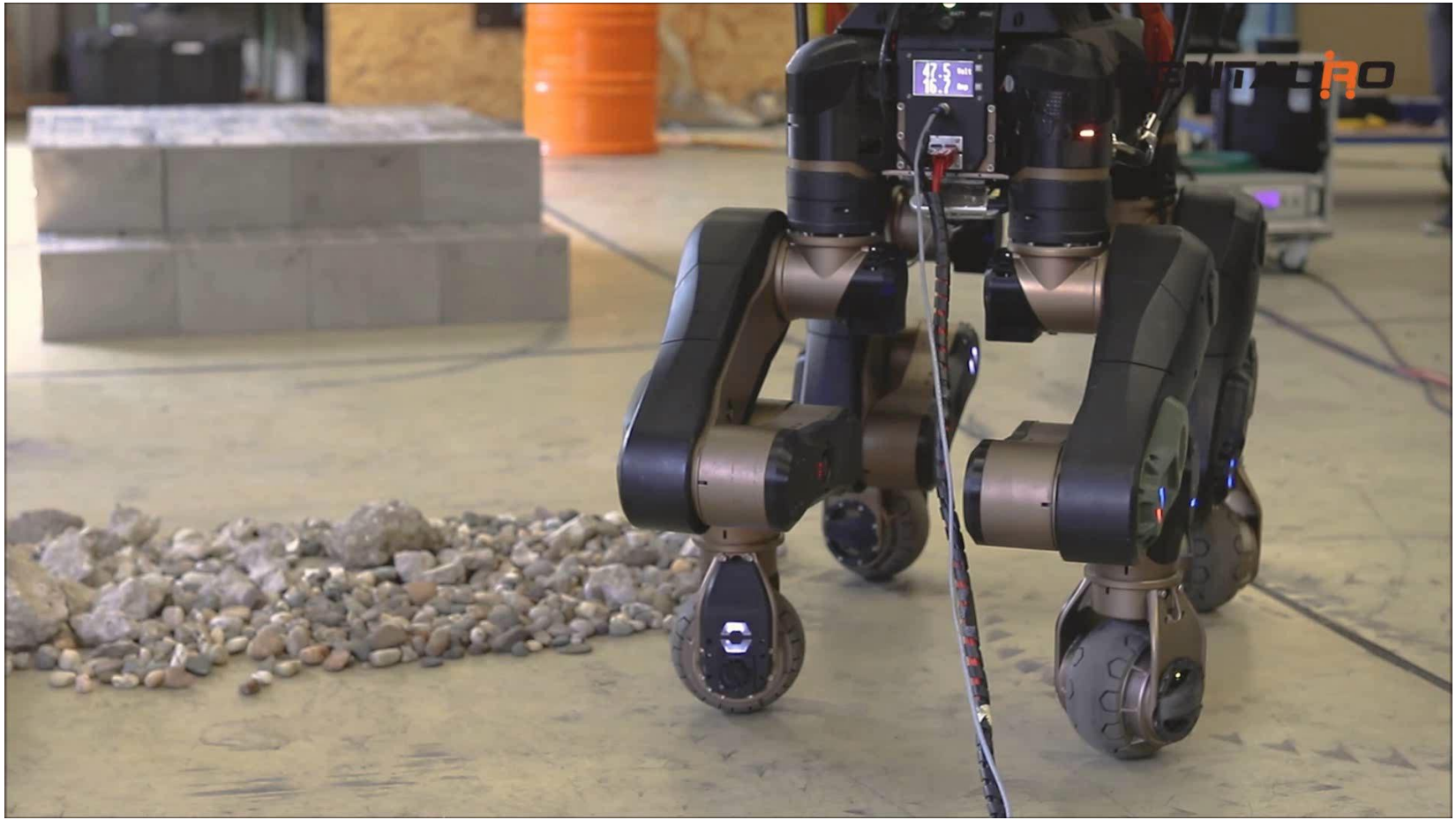
Teleoperation with Joystick and Spacemouse



- Flexible user interfaces for locomotion and manipulation tasks
- 3D situation awareness
- Motion editor

[Klamt et al., Journal of Field Robotics 2020]

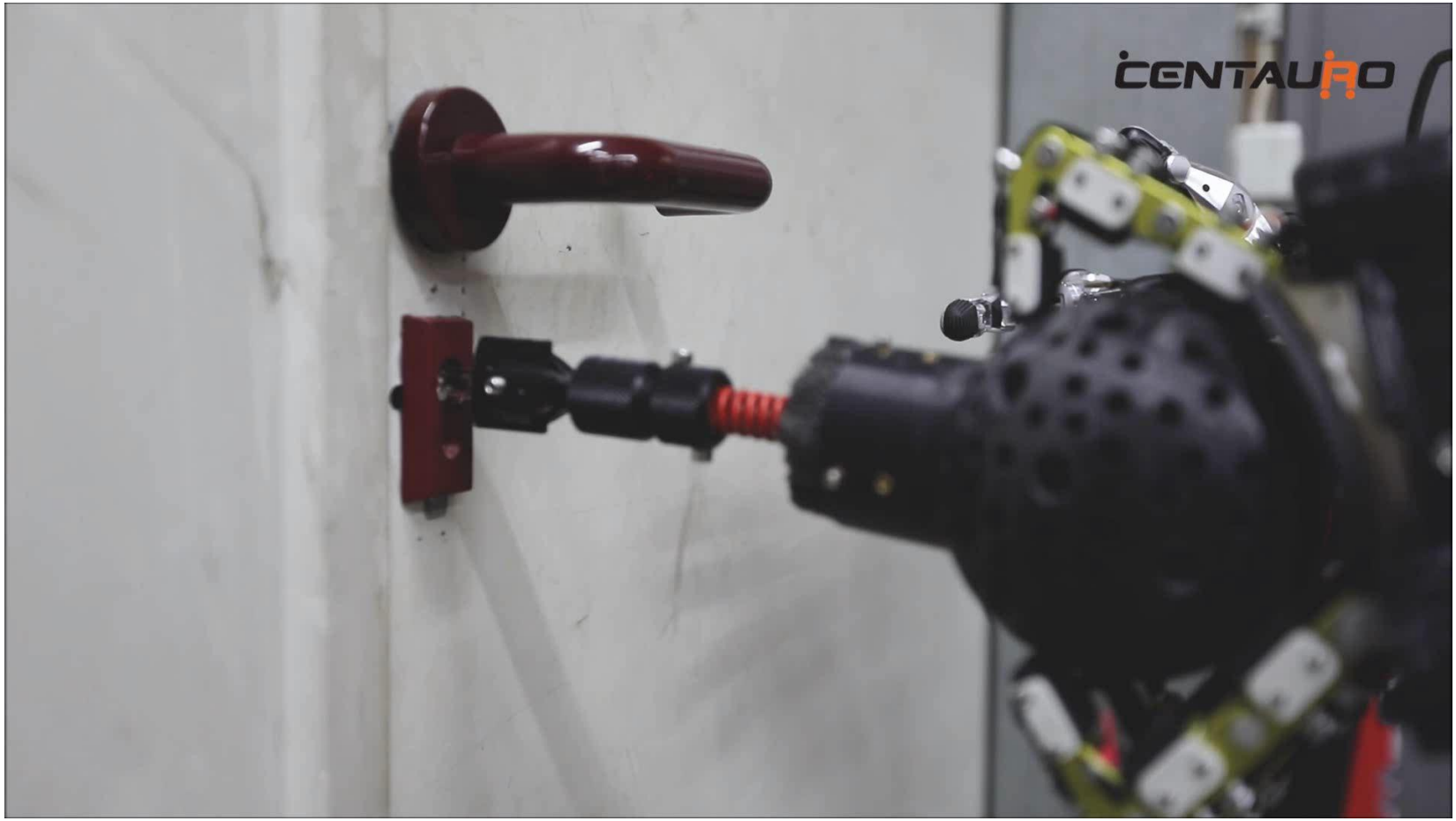
CENTAURO Evaluation @ KHG: Locomotion Tasks



[Klamt et al. RAM 2019]

Grasping an Unknown Power Drill and Fastening Screws

CENTAURO: Complex Manipulation Tasks



[Klamt et al. RAM 2019]

ANA Avatar XPRIZE Competition

- Organized by XPRIZE Foundation
- Sponsored by All Nippon Airways (ANA)
- **Objective**: Create a robotic avatar system that can transport human senses, actions, and presence to a remote location in real time
  - Expanding human connection
  - Transferring skills
  - Exploring dangerous or inaccessible places
- Panel of 22 expert judges
- Launched 03/2018
- **Prize purse of $10M**
- 99 teams registered by 09/2019

[XPRIZE]

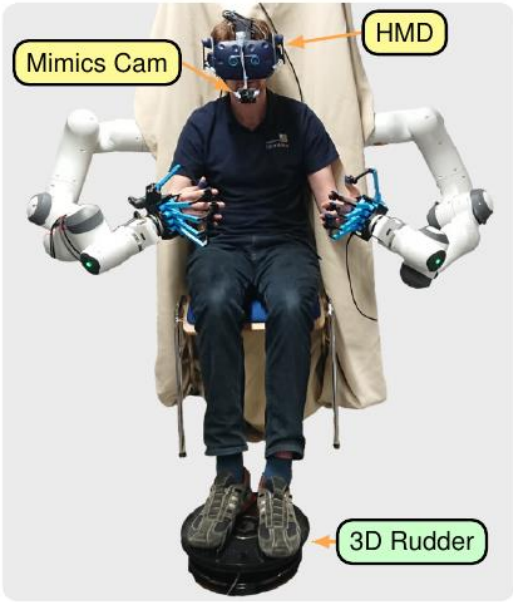
ANA Avatar XPRIZE Competition

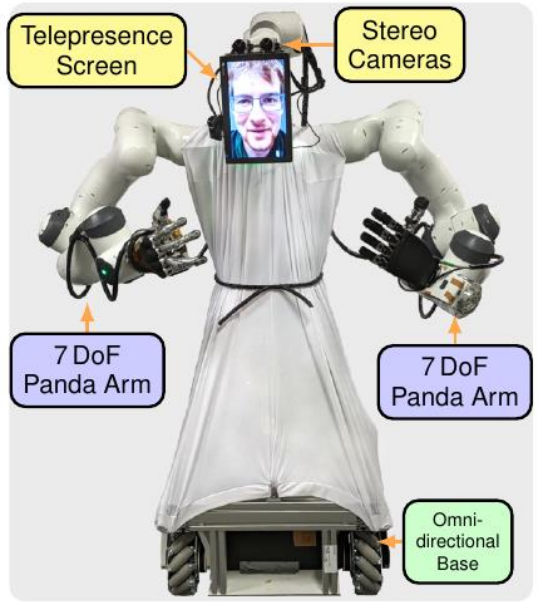
- Required mobility, manipulation, human-human interaction
- Focused on the **immersion** in the remote environment and the **presence** of the remote operator

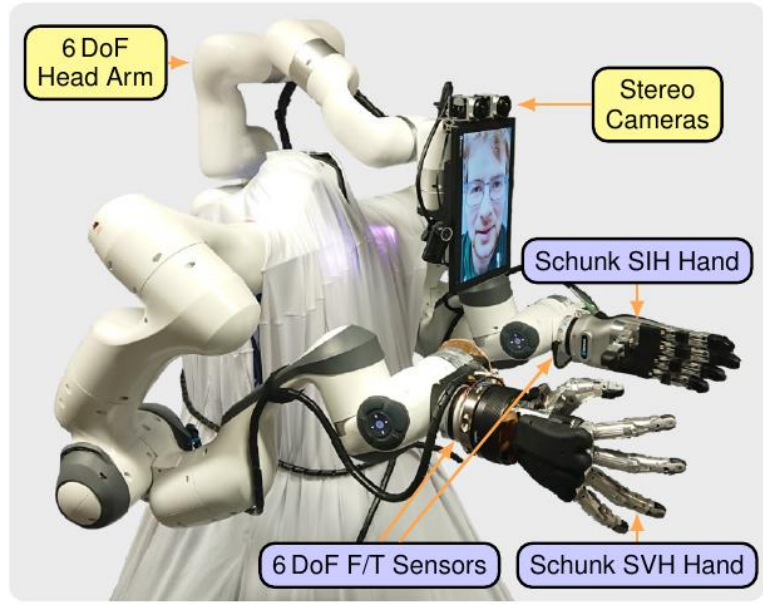
NimbRo Avatar 2021

ANA AVATAR XPRIZE

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



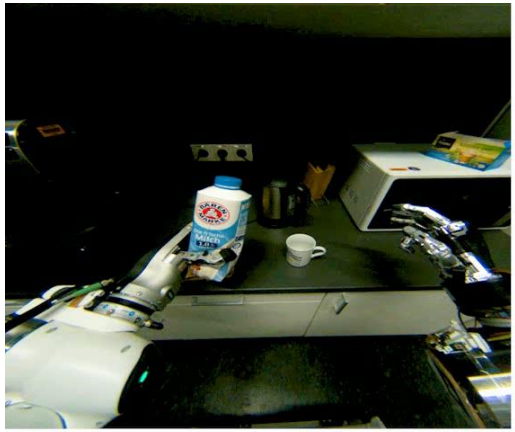




[Schwarz et al. IROS 2021]

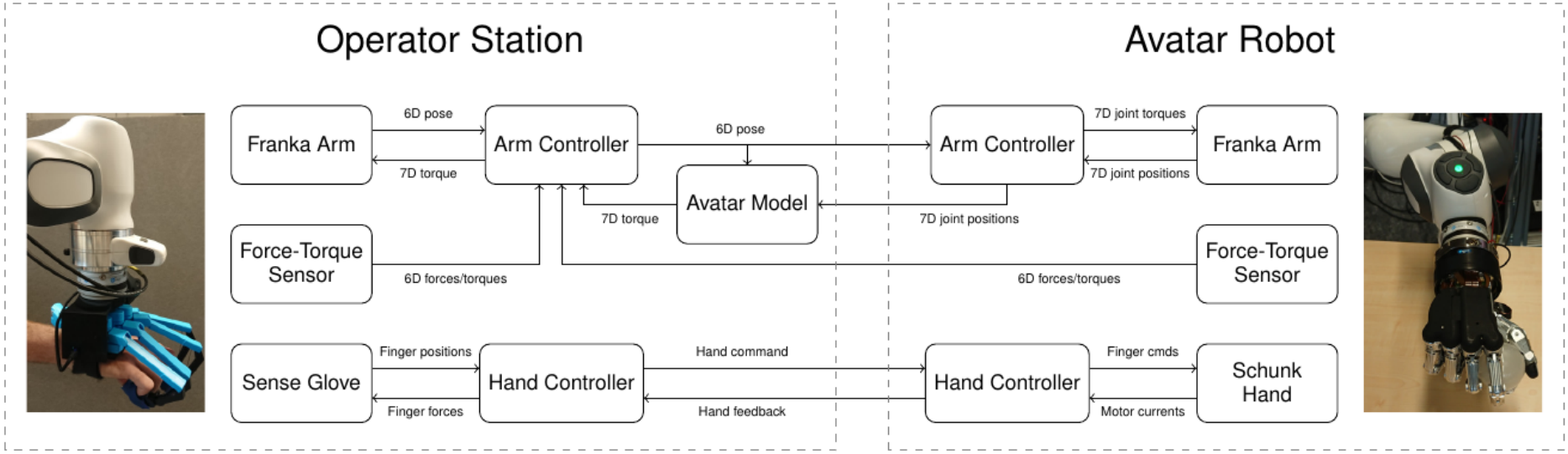
Team NimbRo Semifinal Submission

ANA AVATAR XPRIZE



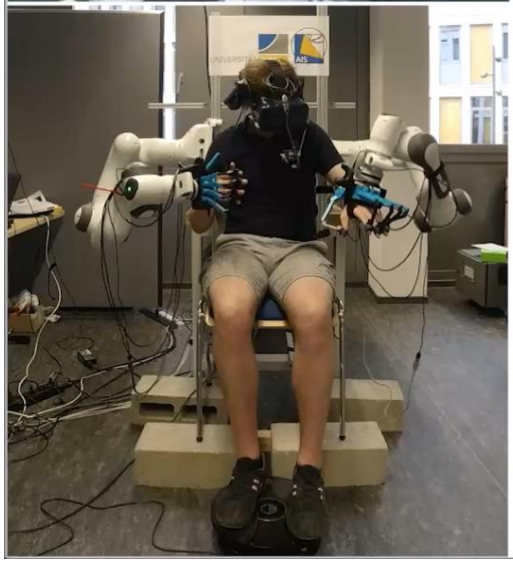
[Schwarz et al. IROS 2021]

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

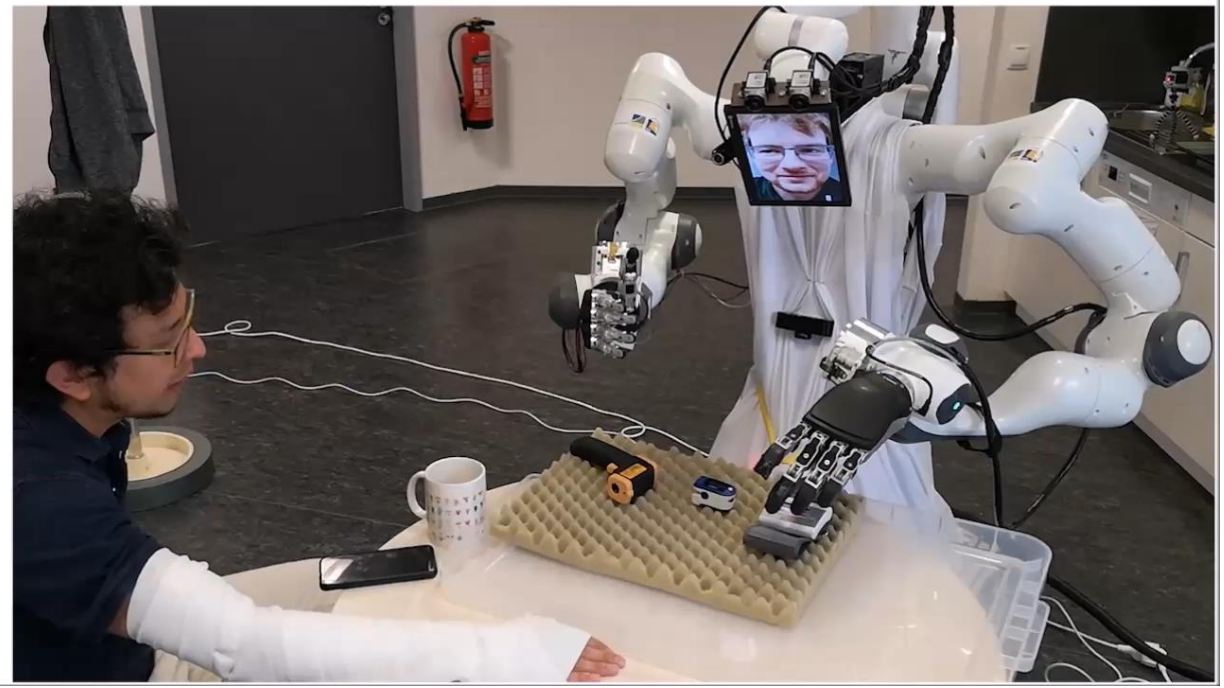
[Lenz and Behnke ECMR 2021]



Team NimbRo Semifinal Team Video

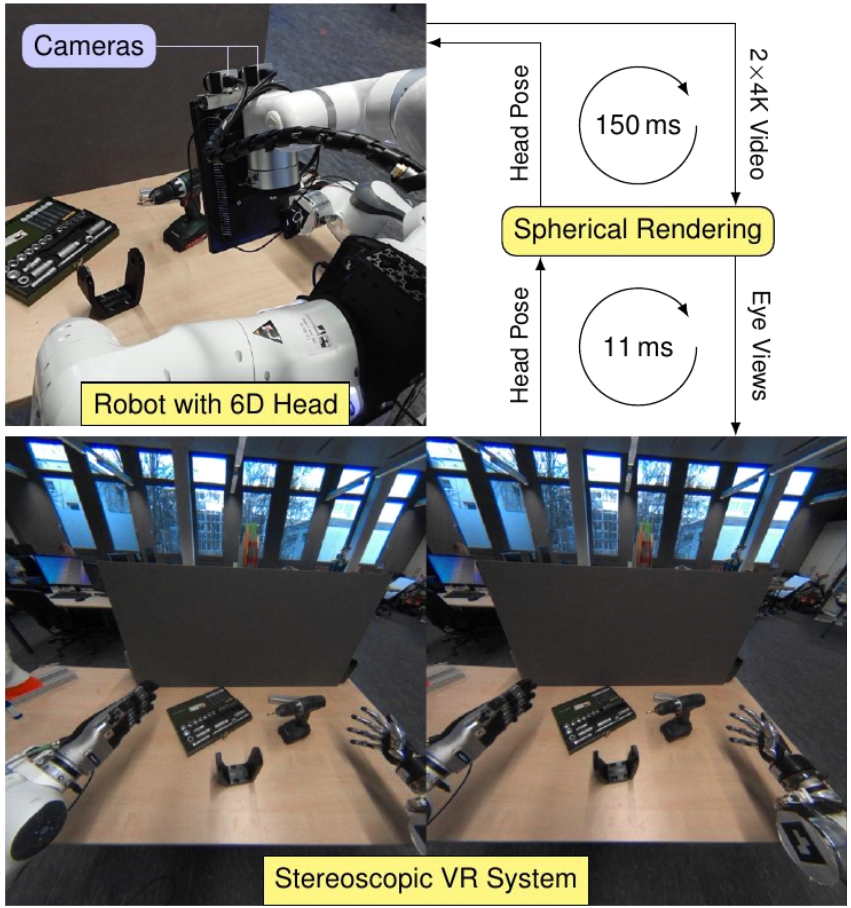
Tasks

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

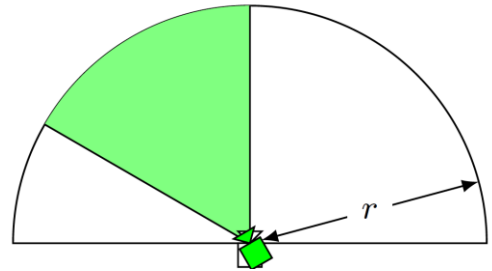


[Schwarz et al. IROS 2021]

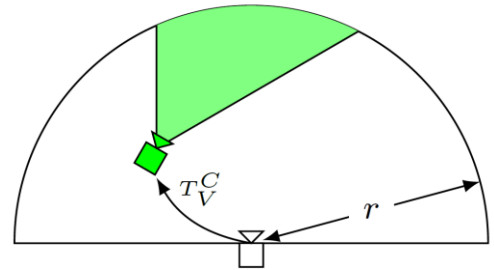
NimbRo Avatar: Immersive Visualization



- 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

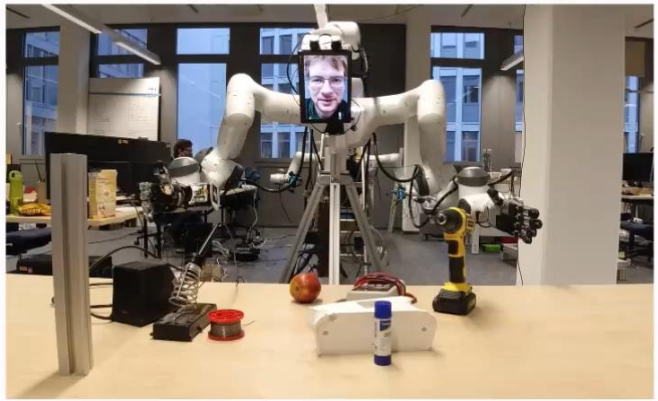


Distortions for translations

[Schwarz and Behnke Humanoids 2021]

NimbRo Avatar: Immersive Visualization

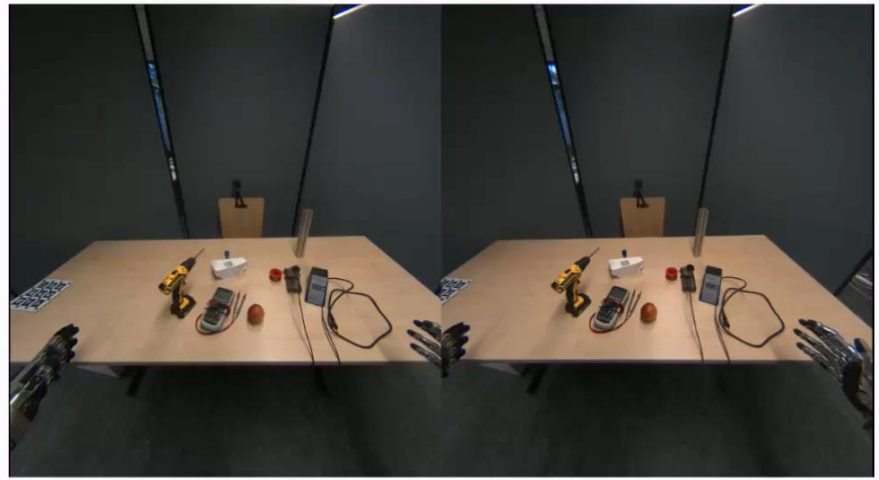
Avatar Robot



Wide-Angle Stereo

Operator

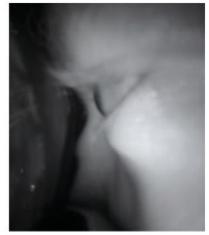
HMD View



[Schwarz and Behnke Humanoids 2021]

NimbRo Avatar: Operator Face Animation

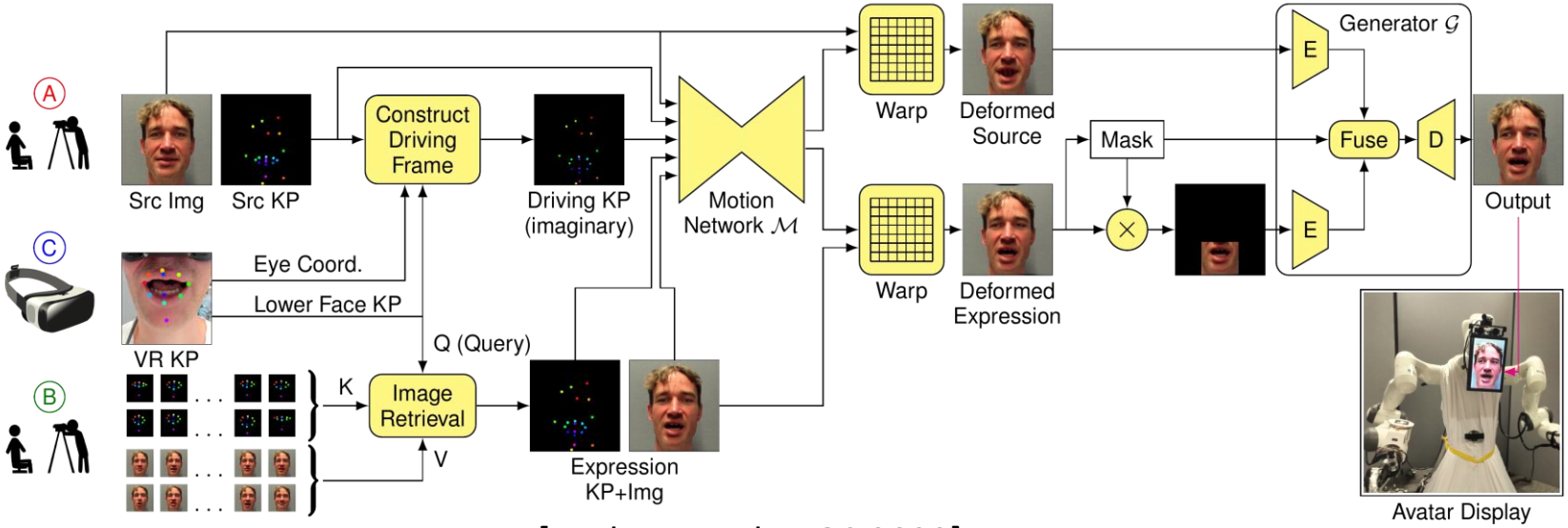
- Operator images without HMD
- Capture mouth and eyes
- Estimate gaze direction and facial keypoints


Left Eye

Mouth

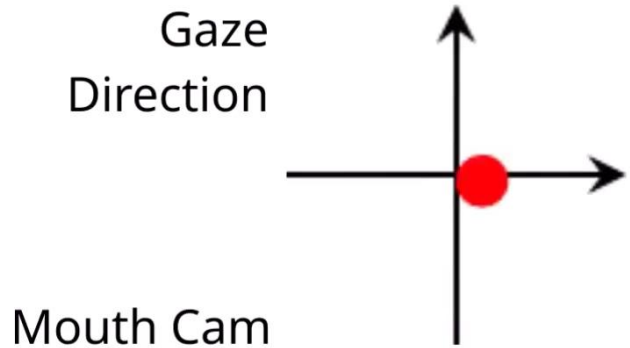
Right Eye

- Generate animated operator face using a warping neural network



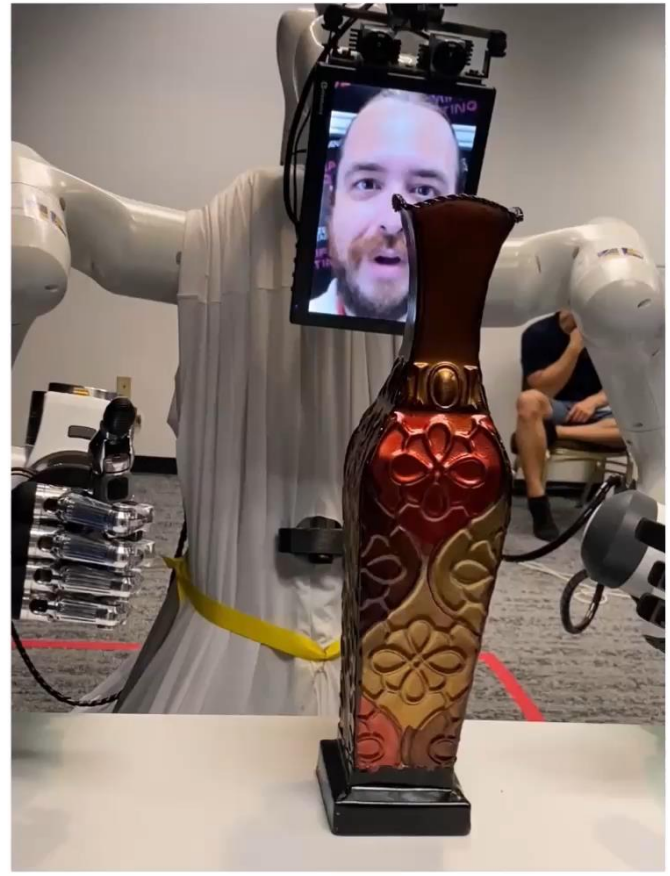
[Rochow et al. IROS 2022]

NimbRo Avatar: Operator Face Animation



Mouth Cam

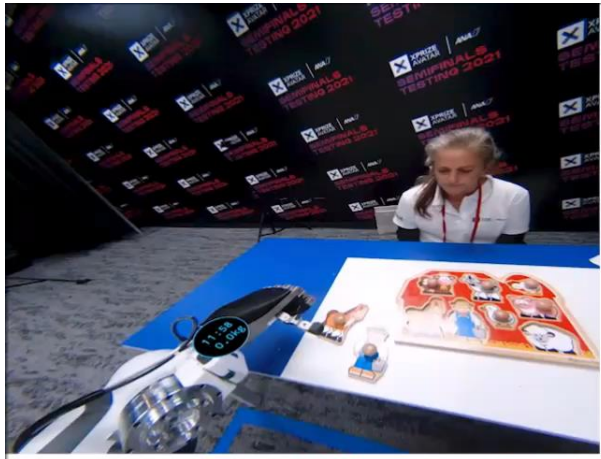
Output

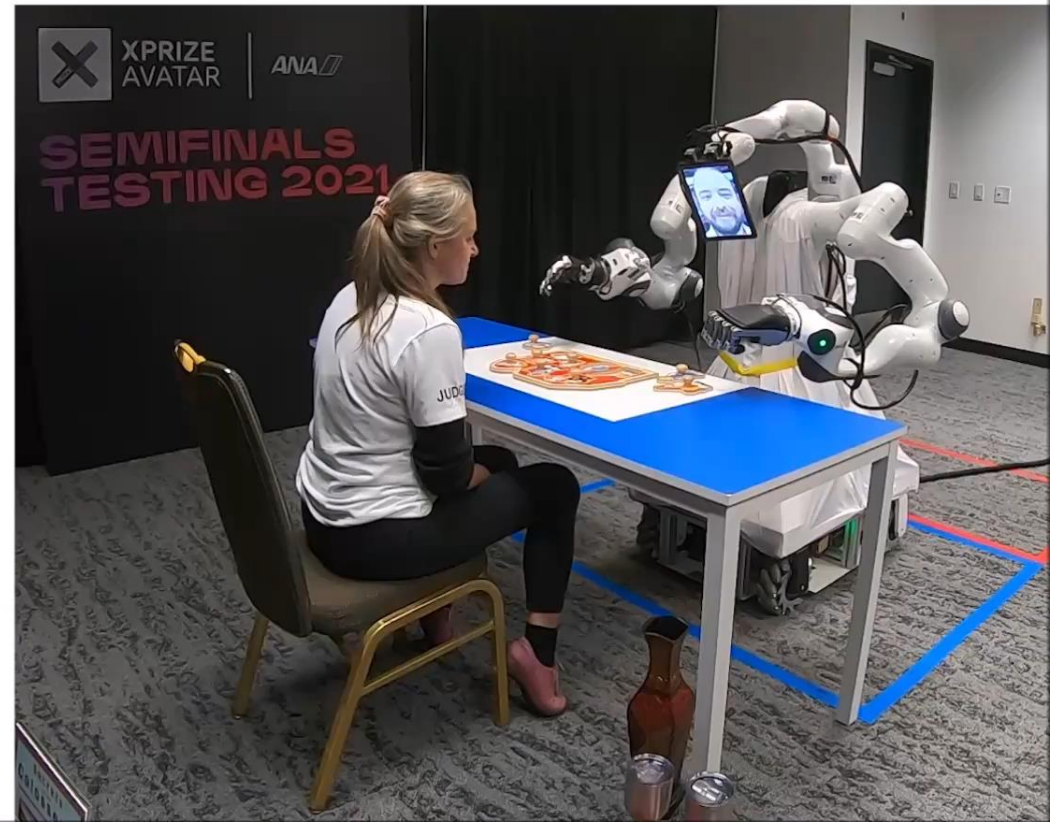


[Rochow et al. IROS 2022]

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

Semifinals Results

Rank	Team Name	Country	Tested in	Score
1	NimbRo	Germany	Miami	99
2	iCub	Italy	own lab	95.25
3	i-Botics	Netherlands	own lab	93.75
4	Team Northeastern	Unites States	Miami	93
5	Dragon Tree Labs	Singapore	Miami	93
6	AVATRINA	United States	Miami	92.75
7	Avatar Hubo	United States	Miami	92
8	Tangible	United States	Miami	92
9	AlterEgo	Italy	own lab	91.75
10	Cyberselves	Un. Kingdom	Miami	90.75
11	Team SNU	South Korea	Miami	89.5
12	Pollen Robotics	France	Miami	89.5
13	Last Mile	Japan	Miami	88.5
14	Enzo	Colombia	own lab	87.25
15	Team UNIST	South Korea	Miami	86
16	Inbiodroid	Mexico	Miami	84.5
17	Rezillient	United States	Miami	84
18	Touchlab	Un. Kingdom	Miami	82.5
19	AvaDynamics	United States	Miami	80.5
20	Janus	France/Japan	own lab	80

[XPRIZE]

New Finals Requirements

ANA AVATAR XPRIZE

- Untethered avatar robot, more mobility
- Movable operator station
- Mission on a distant planet
- 10 tasks must be solved in given sequence
- 11/2022: Qualification day, two testing days with daily down-selection of teams

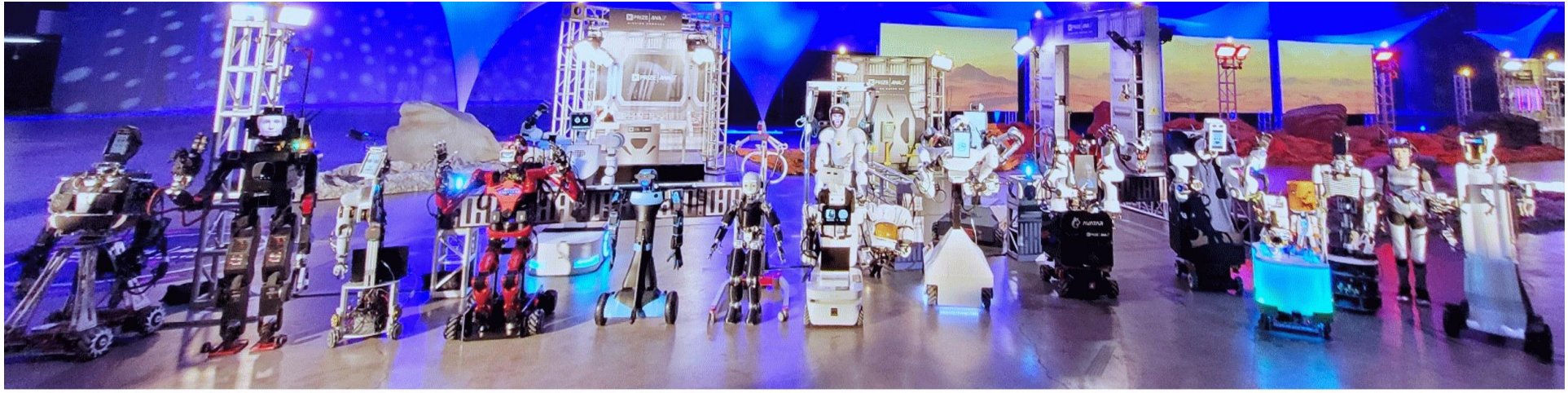
➔ **System reliability extremely important**

Long Beach, CA, USA

Finals Testing Arena

Finals Teams

- 17 teams from 10 countries
- Top research groups and companies



AvaDynamics UNIST i-Botics Tangible AVATRINA Pollen Janus

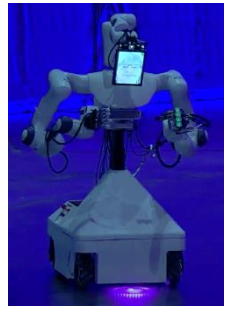
Inbiodroid Avatar-Hubo SNU AlterEgo iCub Cyberselves NimbRo Northeastern Last Mile Dragon Tree Labs

[Behnke et al. Robotics and Automation Magazine 2023]

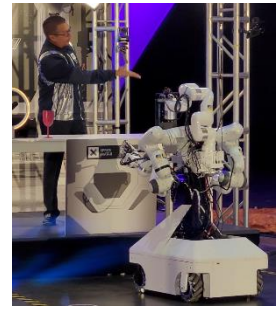
Finals Tasks

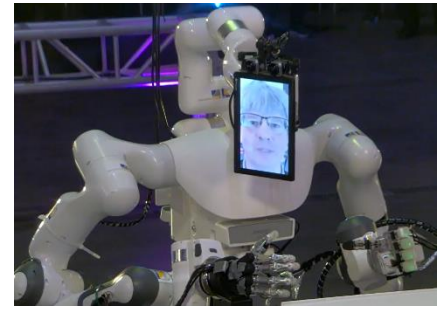
- Three domains:
  - Connectivity
  - Exploration
  - Skill transfer
- Incl. judging object weight and remote feeling of texture
- One point per task
- Tasks fulfillment had highest importance in scoring
- Trial time to break ties

[XPRIZE]

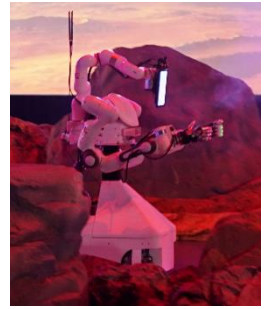

Start

1. Move

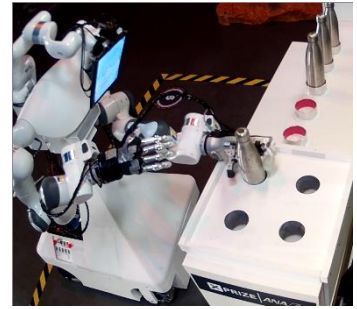

2. Introduce

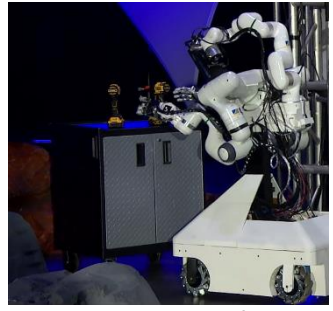

3. Confirm mission

4. Activate switch

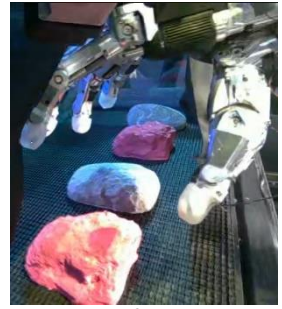

5: Travel planet

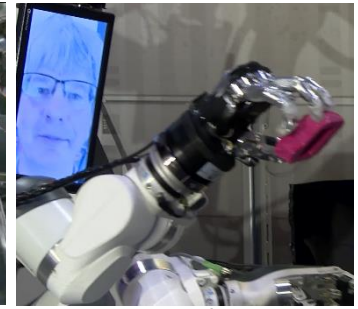
6. Identify full canister


7: Place it


8. Narrow pathway

9: Use drill


10. Feel texture


Finish

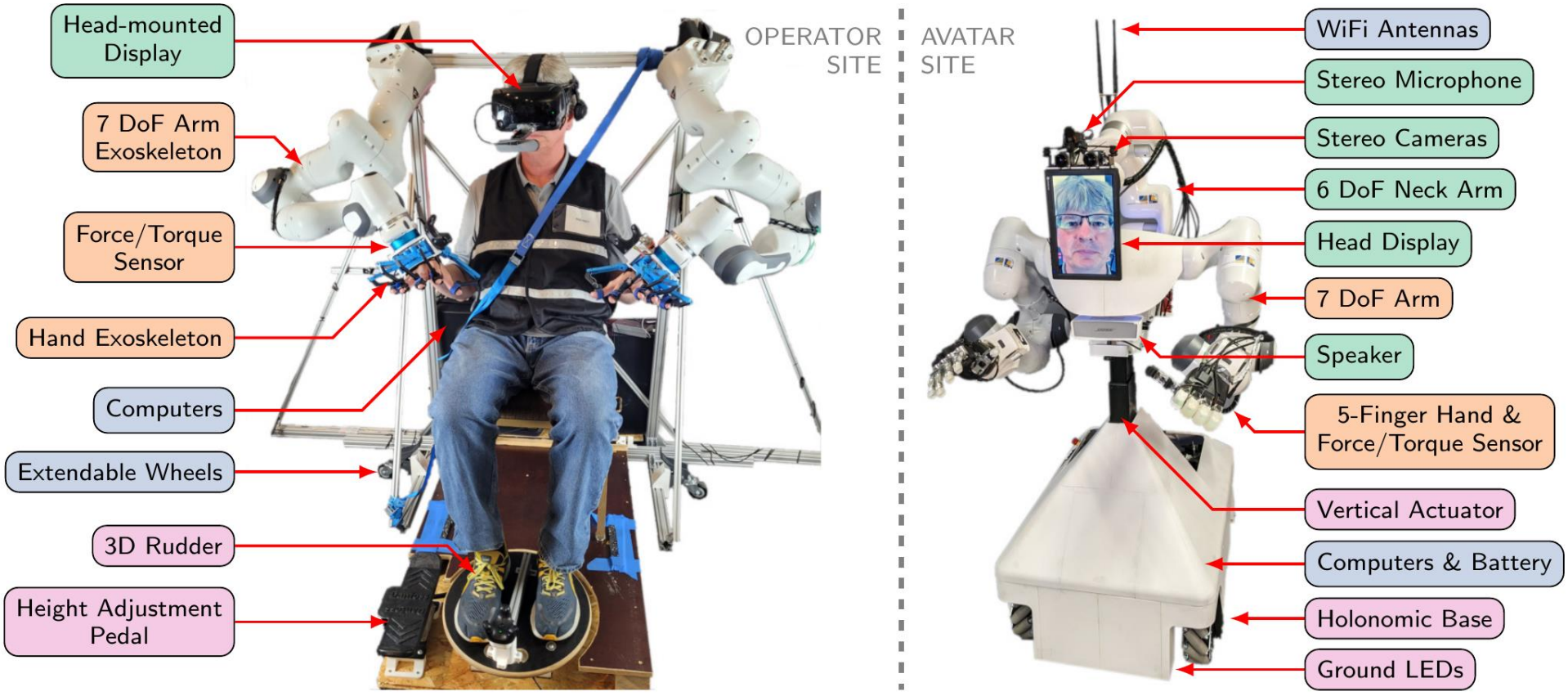
Finals Judged Scoring

- **Operator Experience** (3 points)
  - The avatar system enabled the operator judge to **feel present** in the remote space and conveyed appropriate sensory information.
  - The avatar system enabled the operator judge to **clearly understand** (both see and hear) the recipient.
  - The avatar system was **easy and comfortable** to use.

- **Recipient Experience** (2 points)
  - The avatar robot enabled the recipient judge to feel as though the **remote operator was present** in the space.
  - The avatar robot enabled the recipient judge to **clearly understand** (both see and hear) the operator.

[XPRIZE]

NimbRo Avatar Finals System

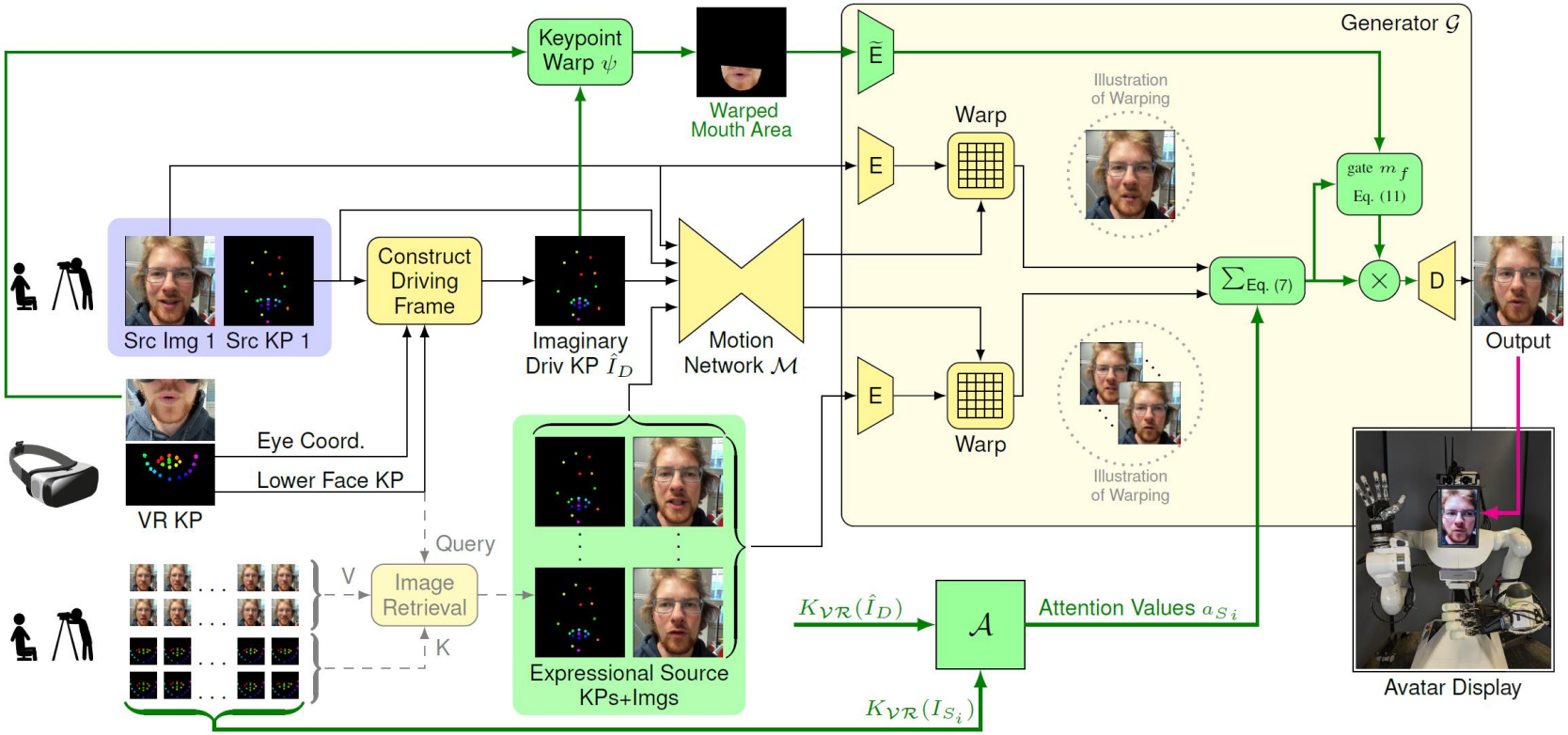


[Lenz et al. International Journal of Social Robotics 2023]

Finals Test Run Day 1

Improved Operator Face Animation

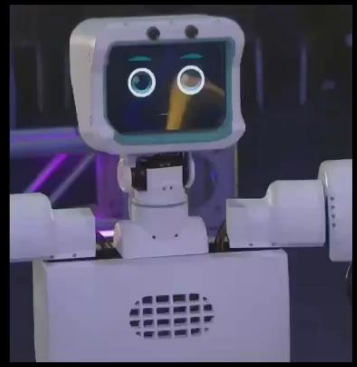
- Direct incorporation of mouth video
- Better temporal continuity



[Rochow et al. IROS 2023]

Face Animation @ Finals

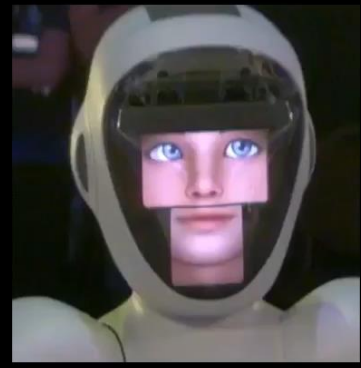
Team UNIST



Ours (NimbRo)

Team AVATRINA [13]

Northeastern [12]



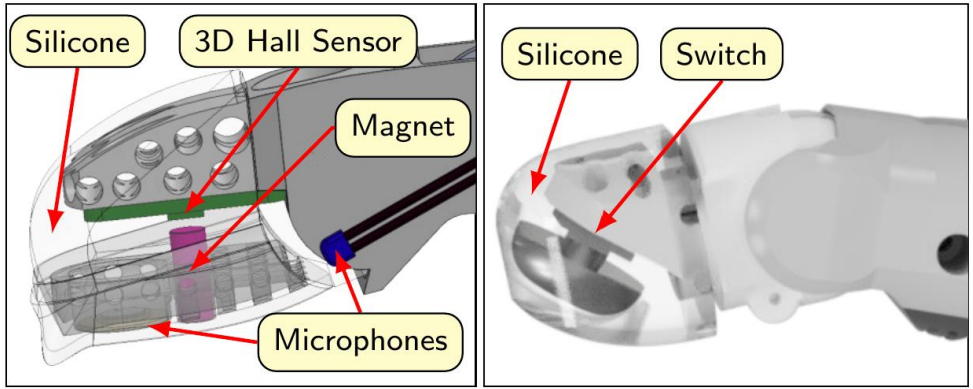
i-BOTICS

Pollen Robotics

Source: Official XPRIZE Avatar live stream

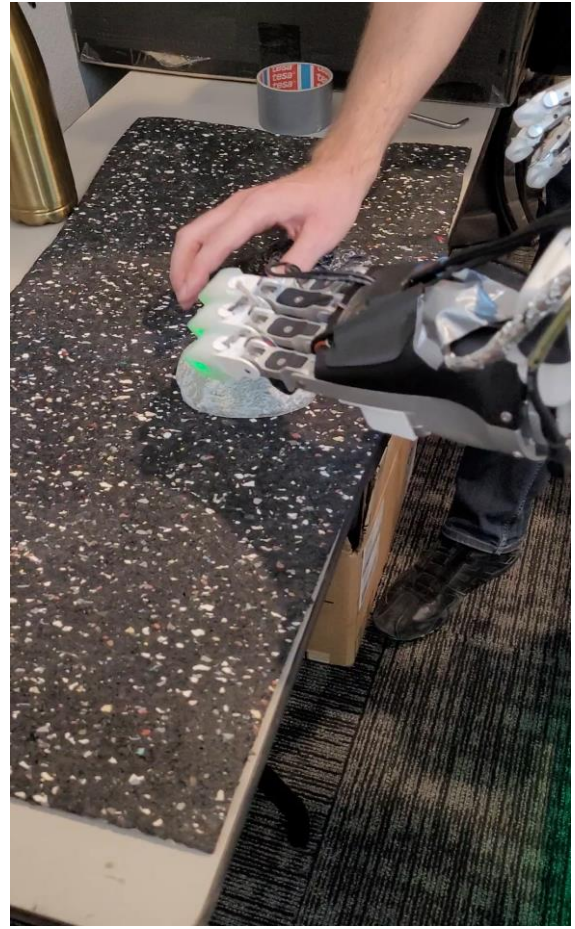
Haptic Perception

- Sensors in the finger tips

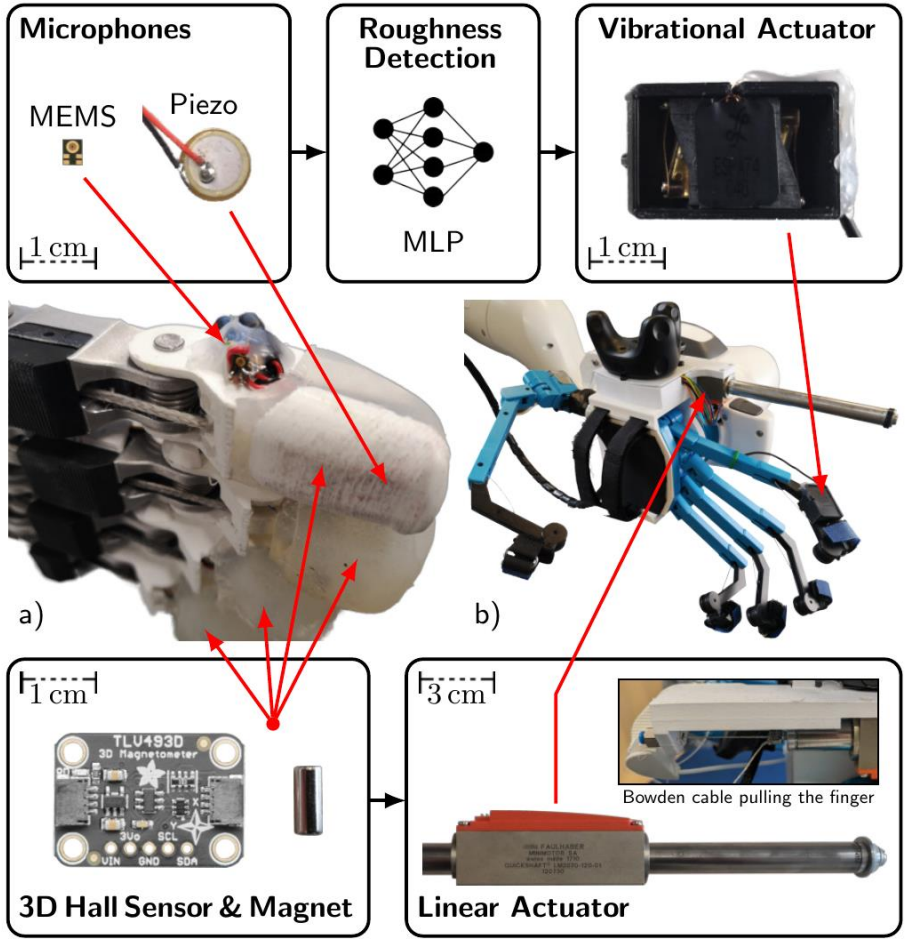


- Actuators on the hand exoskeleton

[Pätzold et al. SMC 2023]



Roughness Perception

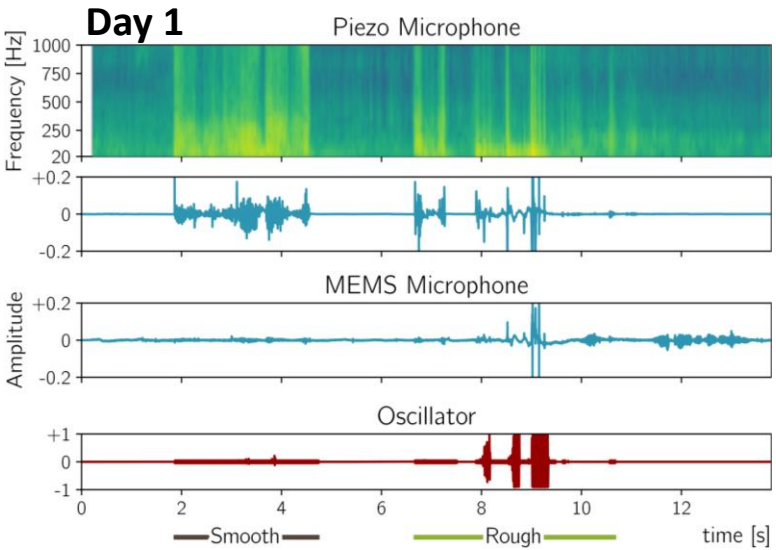


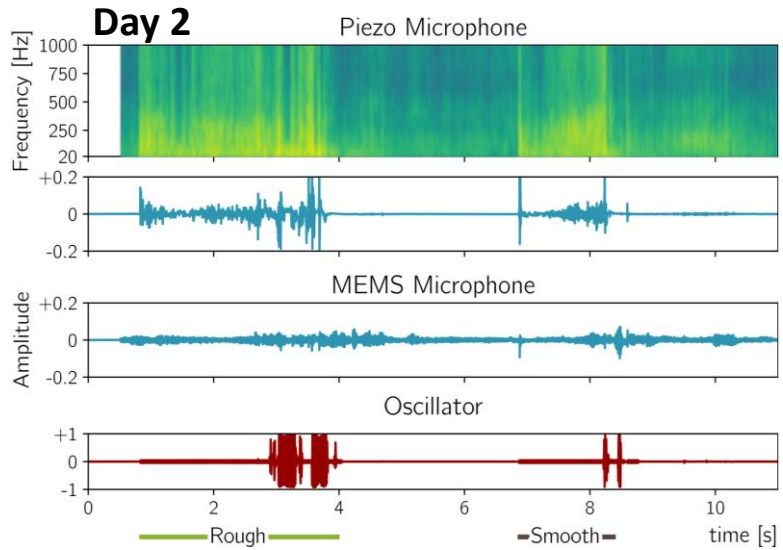
Dataset of Rough and Smooth Objects

[Pätzold et al. SMC 2023]

Finals Task 10: Retrieve a Rough Stone

- Vision partially blocked by a curtain
- 5 stones (3 smooth + 2 rough)





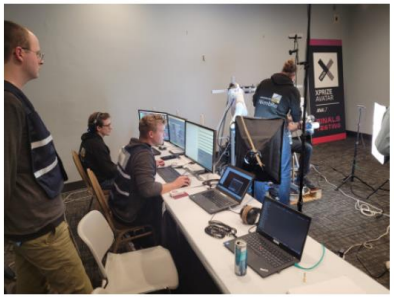
[Pätzold et al. SMC 2023]

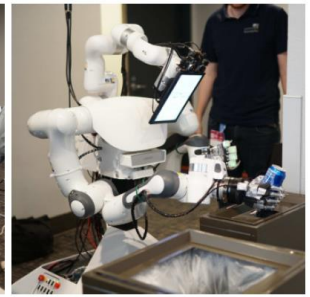
Operator Training

Introduction

Locomotion

Grasping


Monitoring crew

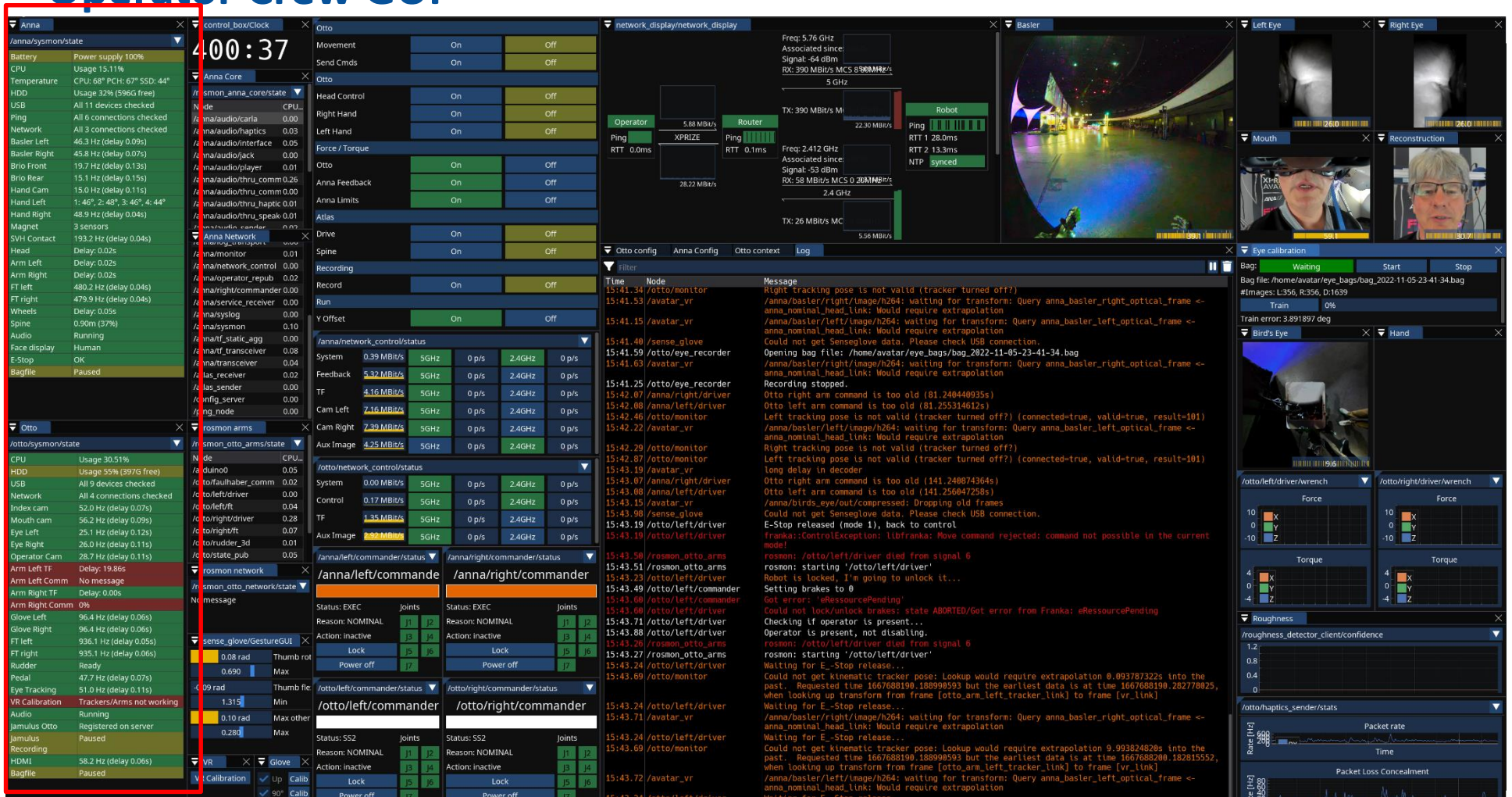

Free experiments

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

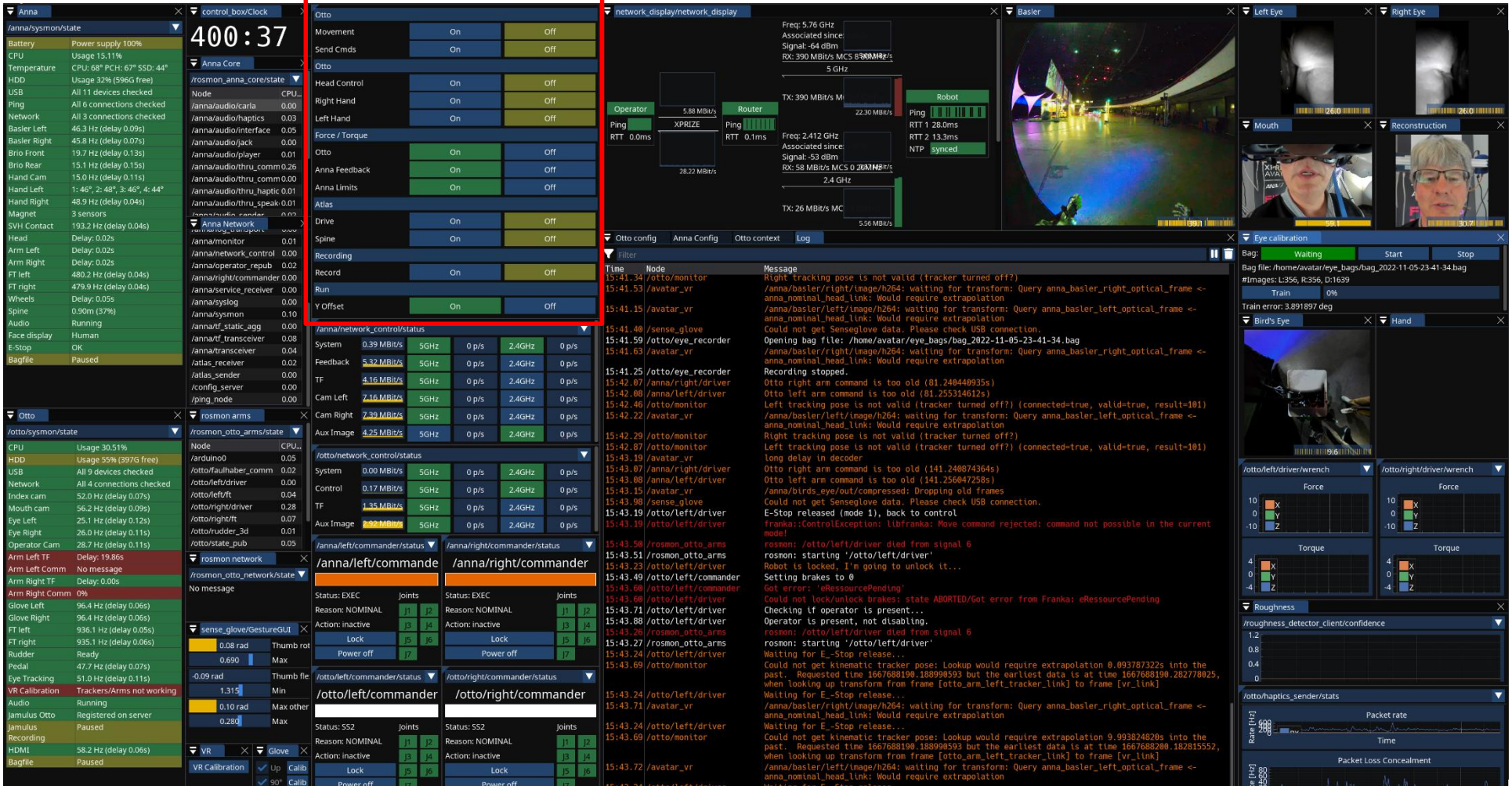
- Dedicated roles: Communication with operator, Software control, Face animation, Hardware support
- Trade-off between learning by own exploration vs. explicit instruction

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

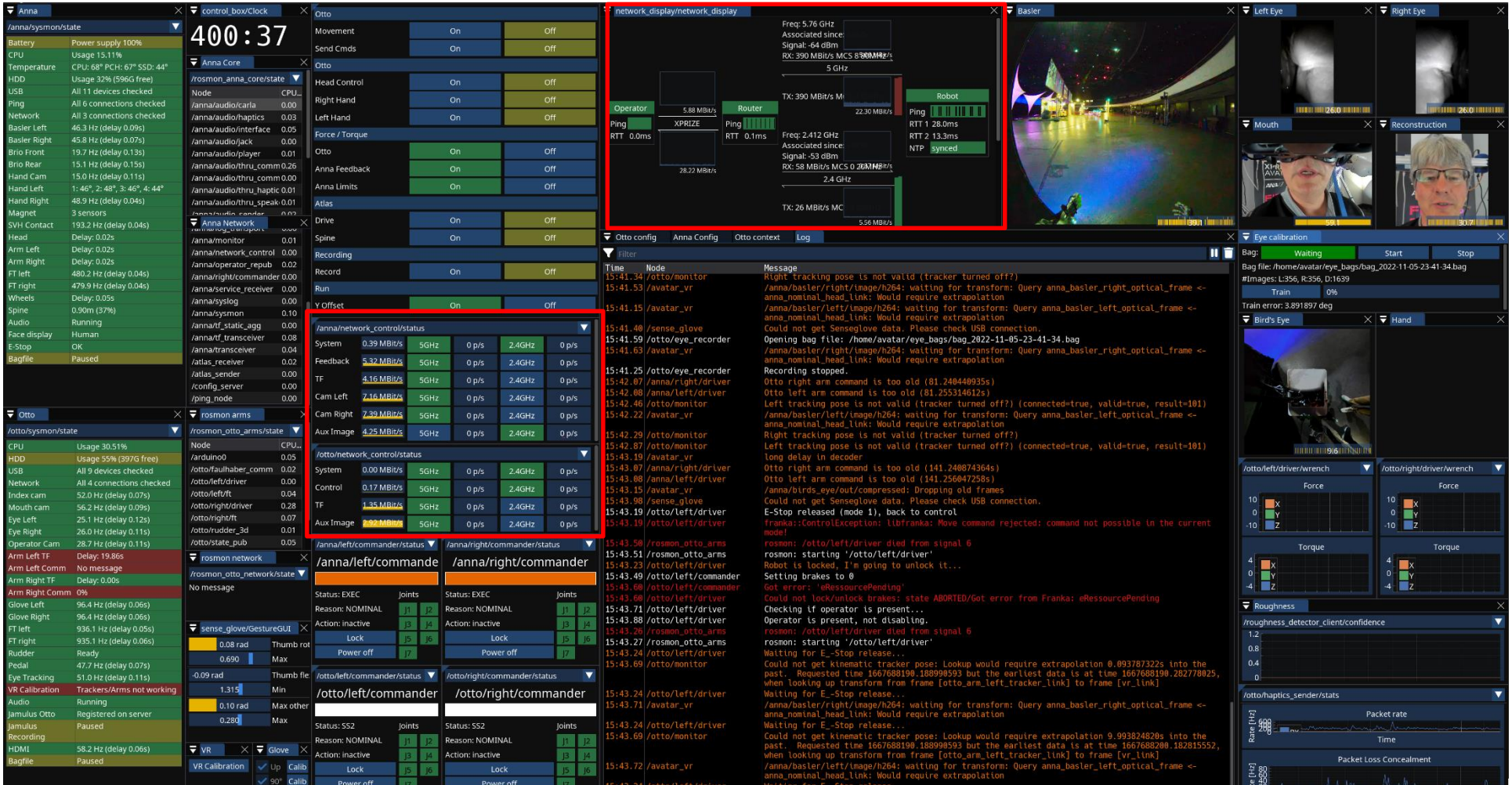
Operator Crew GUI



Operator Crew GUI

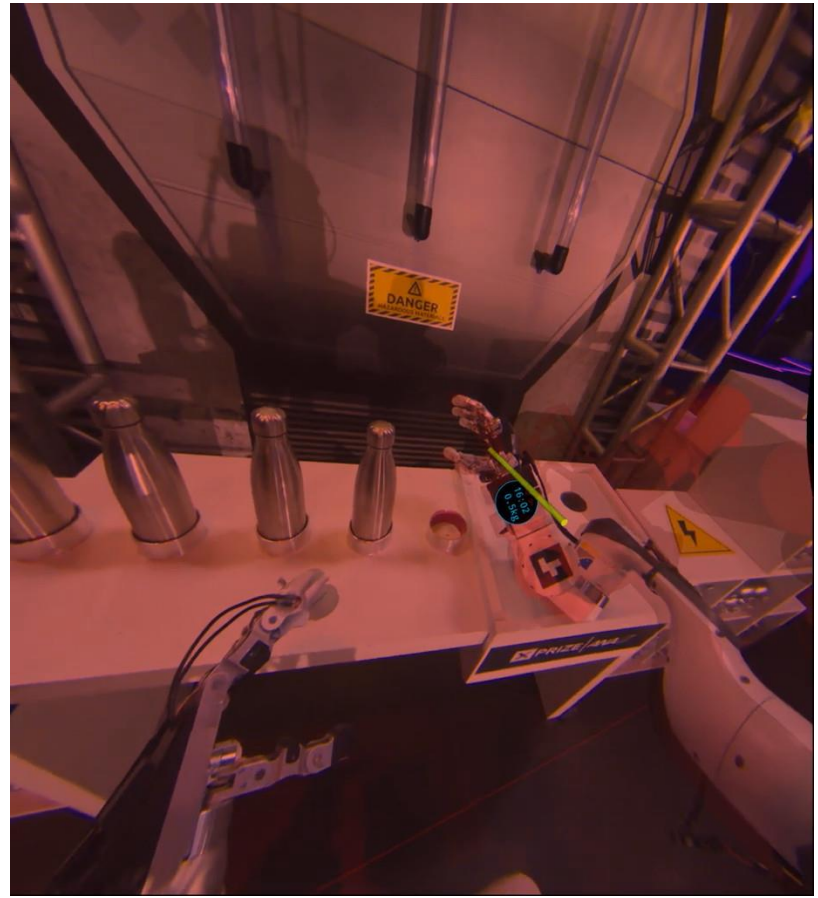


Operator Crew GUI



Reliability Features

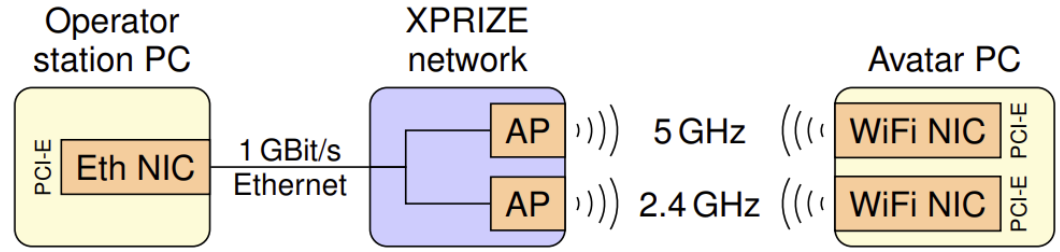
1. Operator crew awareness
2. Automatic arm resets
3. ROS node respawn
4. State- and connectionless network system (pure UDP)
5. Redundant WiFi connections
6. PC watchdog



[Schwarz et al. Humanoids 2023]

Network Details

- Separate ROS cores for operator station and 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
- Monitoring packet loss
- The core software is already open source, more to come:
  https://github.com/AIS-Bonn/nimbro_network



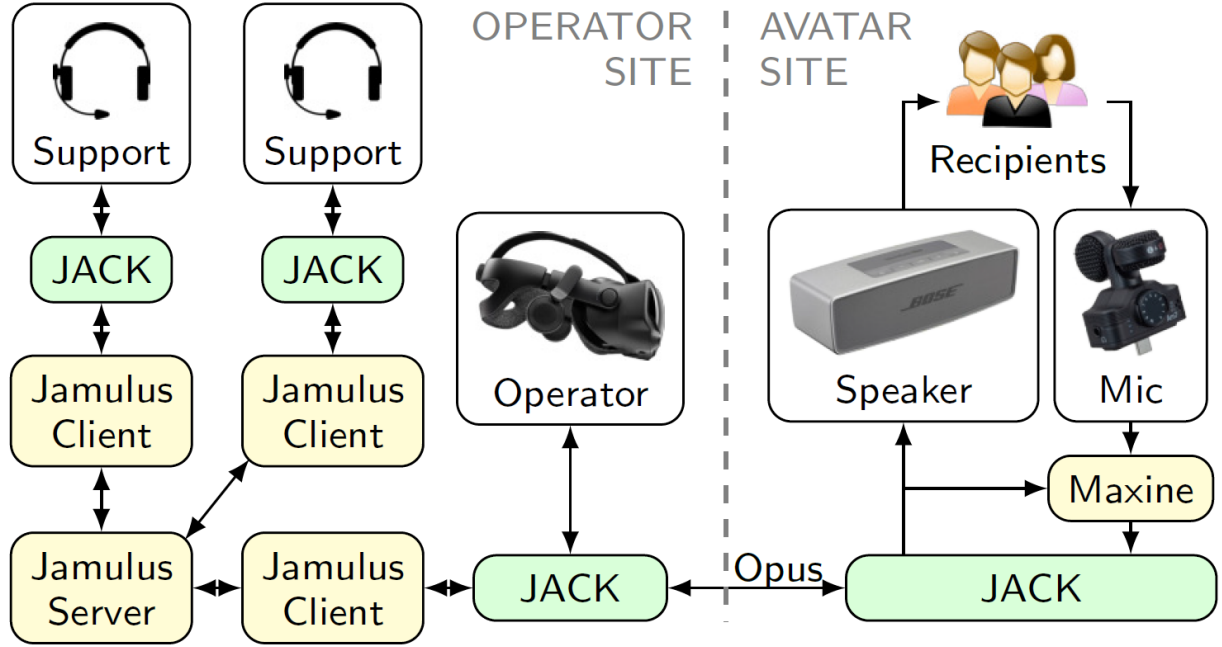
WiFi Bandwidth Requirements

Downlink from avatar				Uplink to avatar			
Channel	MBit/s	5 GHz	2.4 GHz	Channel	MBit/s	5 GHz	2.4 GHz
Arm feedback	8.5	✓	×	Arm control	4.9	✓	✓
Transformations	4.1	✓	×	Transformations	1.4	✓	×
Main cameras	14.7	✓	×	Operator face	5.7	×	✓
Hand camera	5.5	×	✓	Audio	0.4	✓	✓
Diagnostics	0.4	✓	✓				
Audio	0.4	✓	✓				
Total [MBit/s]		28.1	6.3	Total [MBit/s]		6.7	11.0

[Schwarz et al. Humanoids 2023]

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



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

Finals Day 2 Testing

Finals Results

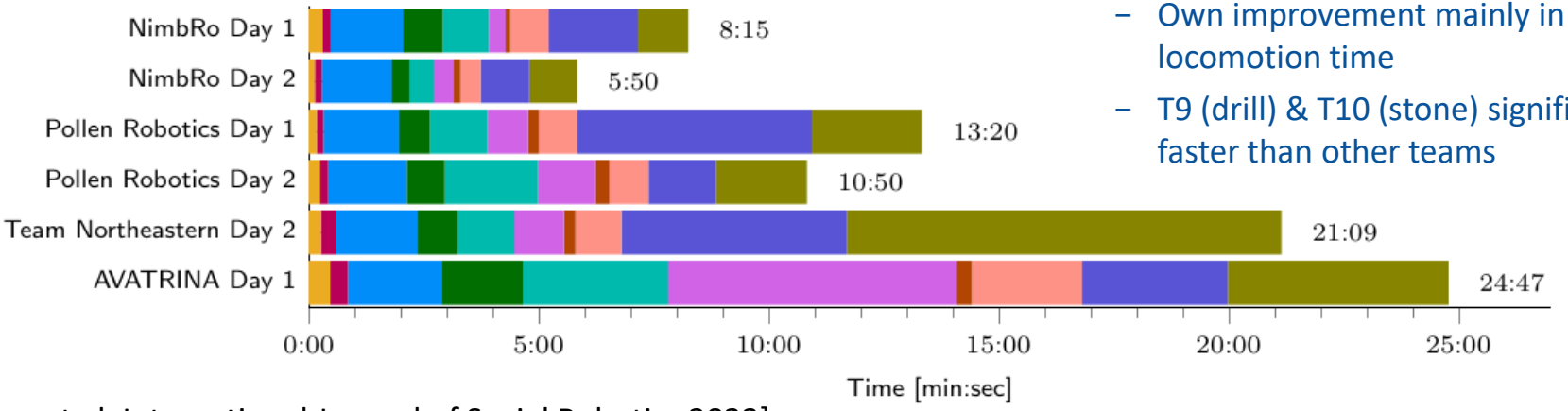
Rank	Team name	Time	Task score	Judged score	Total
1	NimbRo (DE)	5:50	10	5	15
2	Pollen Robotics (FR)	10:50	10	5	15
3	Team Northeastern (US)	21:09	10	4.5	14.5
4	AVATRINA (US)	24:47	10	4.5	14.5
5	i-Botics (NL)	25:00	9	5	14
6	Team UNIST (KR)	25:00	9	4.5	13.5
7	Inbiodroid (MX)	25:00	8	5	13
8	Team SNU (KR)	25:00	8	4.5	12.5
9	AlterEgo (IT)	25:00	8	4.5	12.5
10	Dragon Tree Labs (SG)	25:00	7	4	11
11	Avatar Hubo (US)	25:00	6	3.5	9.5
12	Last Mile (JP)	25:00	5	4	9

[XPRIZE]

Team NimbRo

Finals Timings

Team	Day	Time[1] [mm:ss] Start[2]	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 [25]	1	00:00	00:33	00:24	02:08	01:43	04:03	01:27	00:36	01:56			12:50
		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 [26]	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 [51]	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

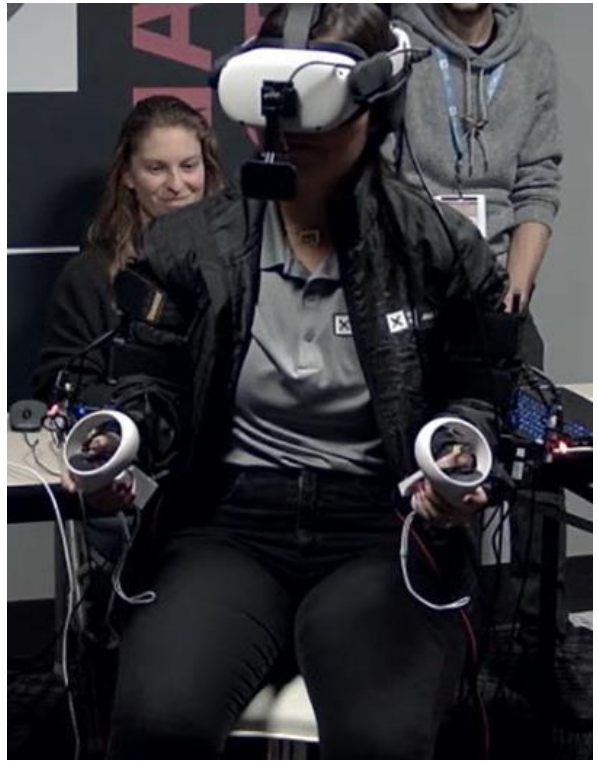


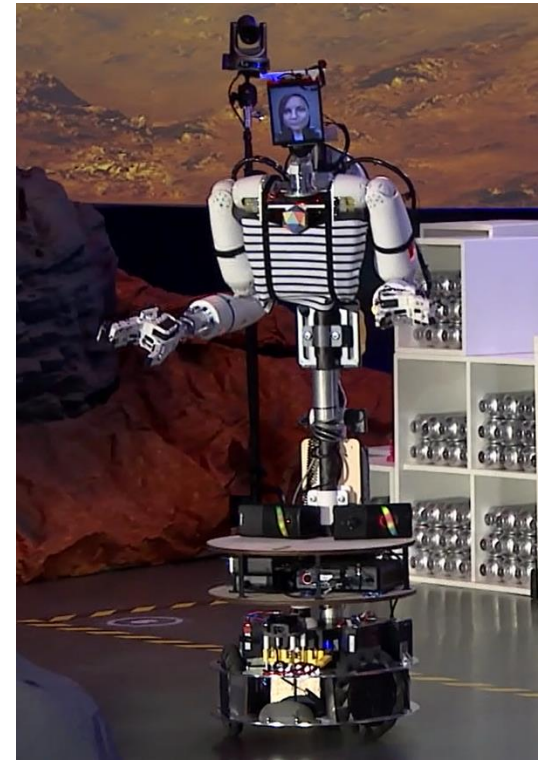
- Own improvement mainly in locomotion time
- T9 (drill) & T10 (stone) significantly faster than other teams

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

2nd Place: Pollen Robotics

- Cost-effective design of robot and operator station
- Human-like upper body with Orbita 3 DoF actuators in shoulder, wrist, and neck

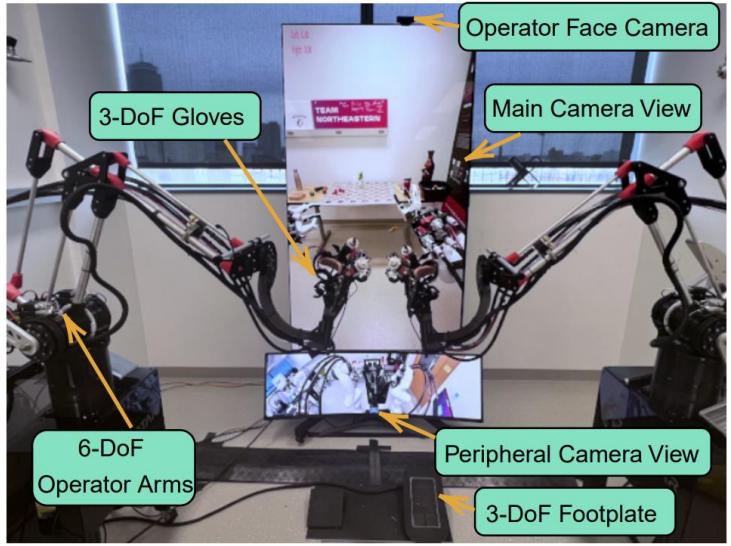


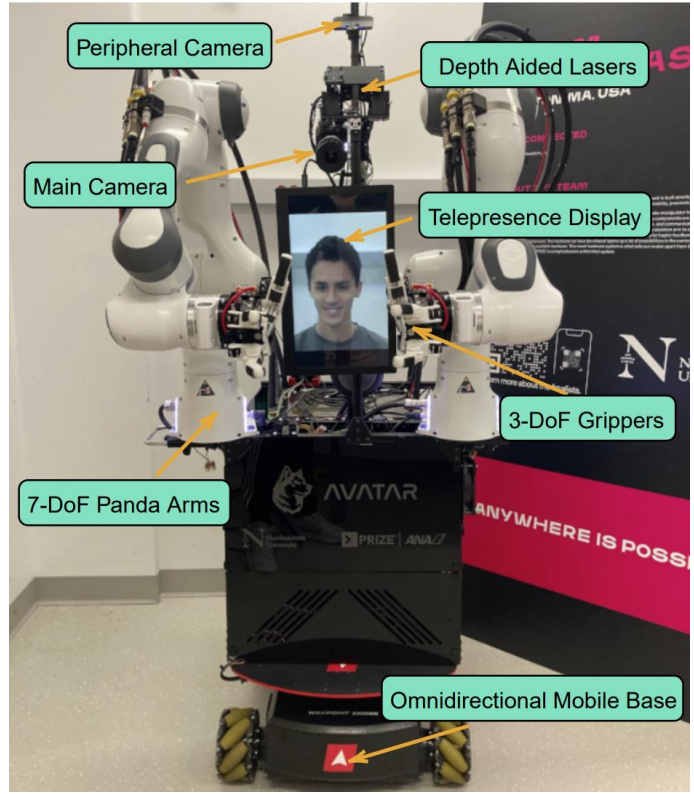


[Pollen Robotics]

3rd Place: Team Northeastern

- Hydraulically actuated glove-gripper pair for haptic force feedback
- Non-immersive visualization with two monitors
- Projected laser lines aid 3D perception

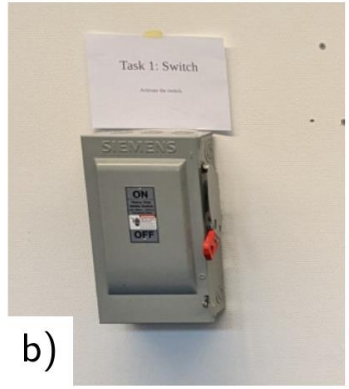




[Luo et al. arXiv:2303.04932]

After the Competition: User Study in our Lab

a)



b)

c)

d)

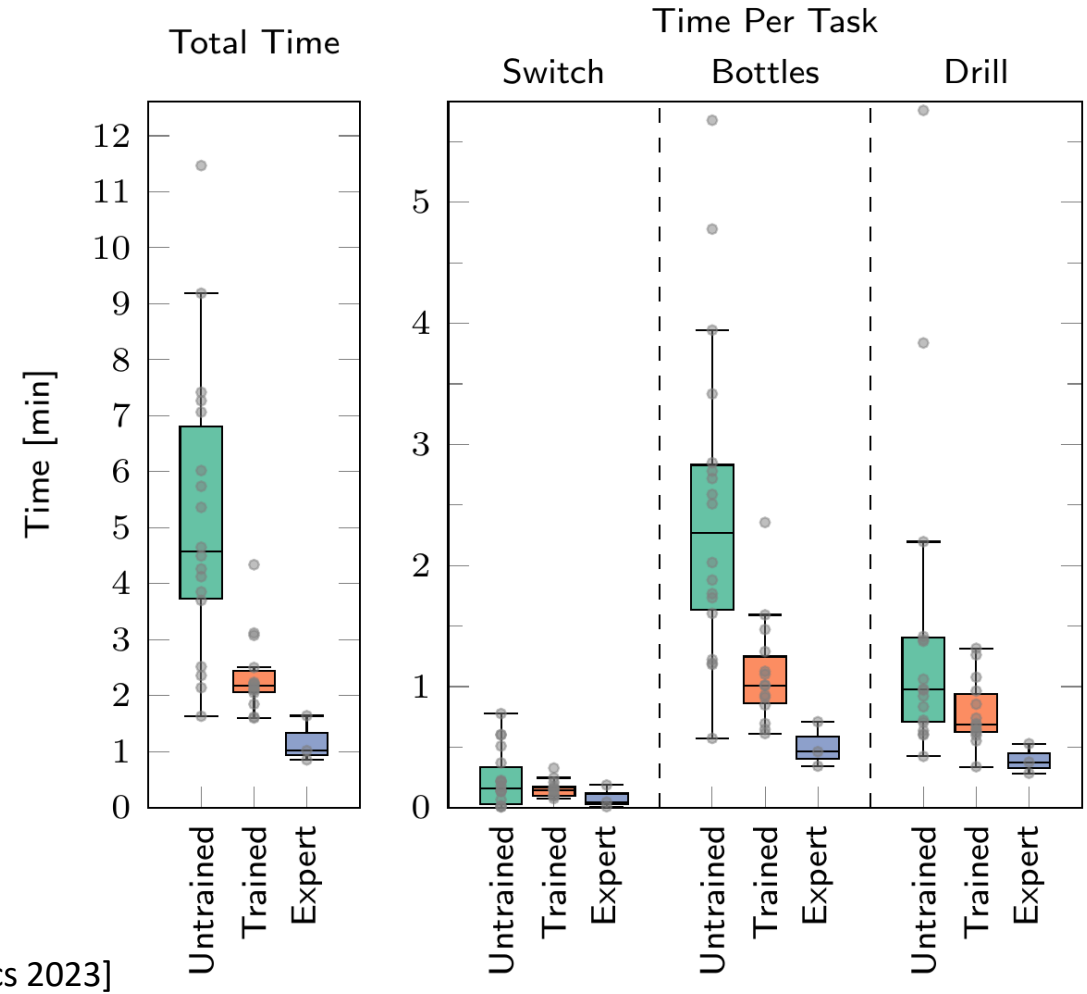
- Three tasks, similar to finals: a), b) Locomotion+Switch, c) Bottles, d) Drill
- 35 participants, 32 with no prior experience of the system
- All participants: 2 min intro video explaining the system (task agnostic)
- Three groups:
  - No training
  - 10 min task training, similar to Finals
  - Expert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

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

After the Competition: User Study in our Lab

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

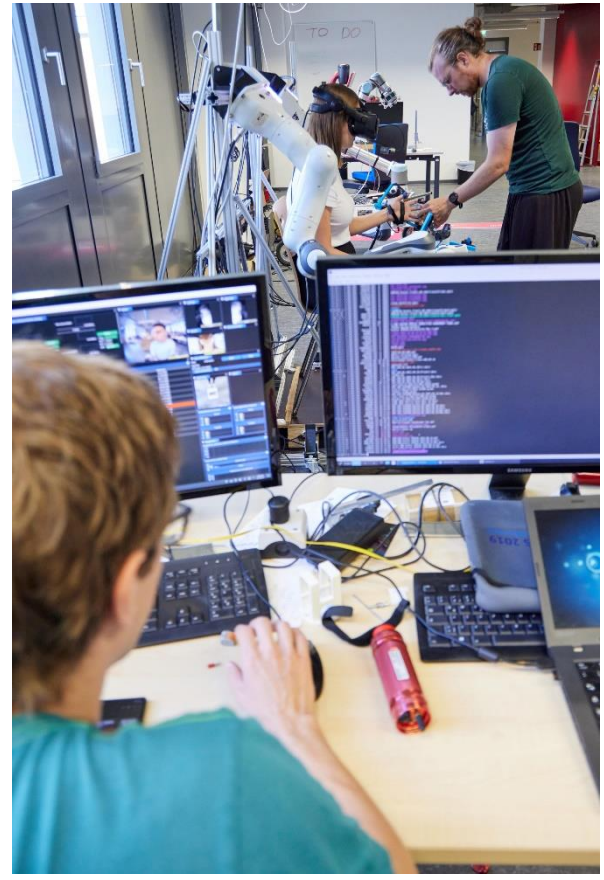
➔ System is very intuitive, but short instruction on tasks improves completion time.



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

Lessons Learned

- Robustness is key
- Latency is the enemy of direct teleoperation
- Frequent testing under competition conditions is essential: System & team!
- 1:1 correspondence is best
- 6D head motion simplifies manipulation control
- Sparse immersive control overlays – don’t break immersion!
- Facial animation and gestures: Head & gaze direction enables shared awareness
- Had to modify components



[Photographer: Volker Lannert]

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

What is Next?

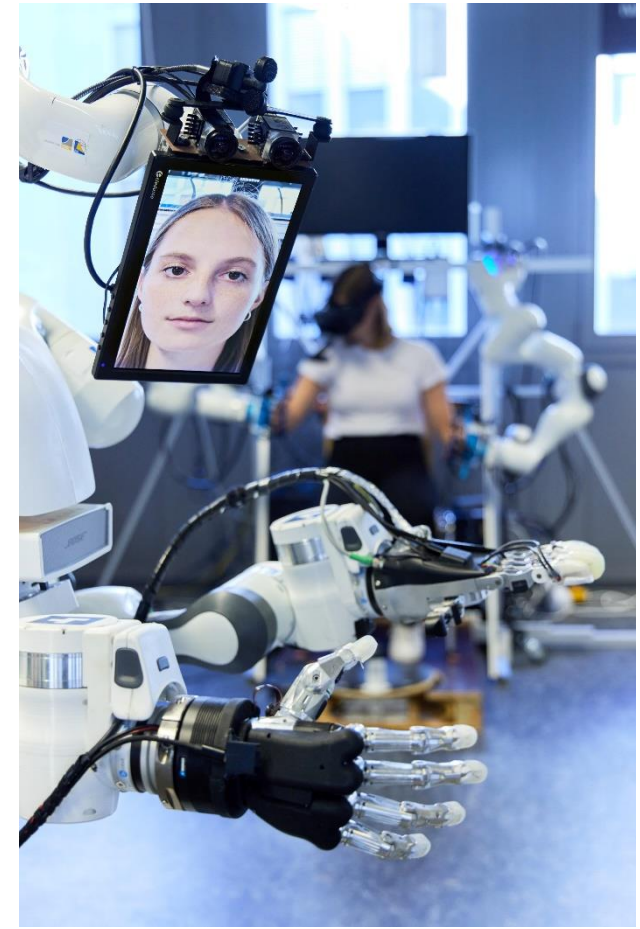
- Transfer to **real applications**
  - Complex avatar systems could be further developed e.g. for
    - Dangerous or hard-to-reach domains,
    - Disaster relief,
    - Medical assistance in isolation wards
  - Everyday virtual travel requires simpler and more affordable systems

[Photographer: Volker Lannert]

- **Research questions** include
  - How much human-likeness avatars should assume?
  - How to address latencies and bandwidth limitations?
  - How to balance and interface direct control and autonomy?

Motivation for Autonomy

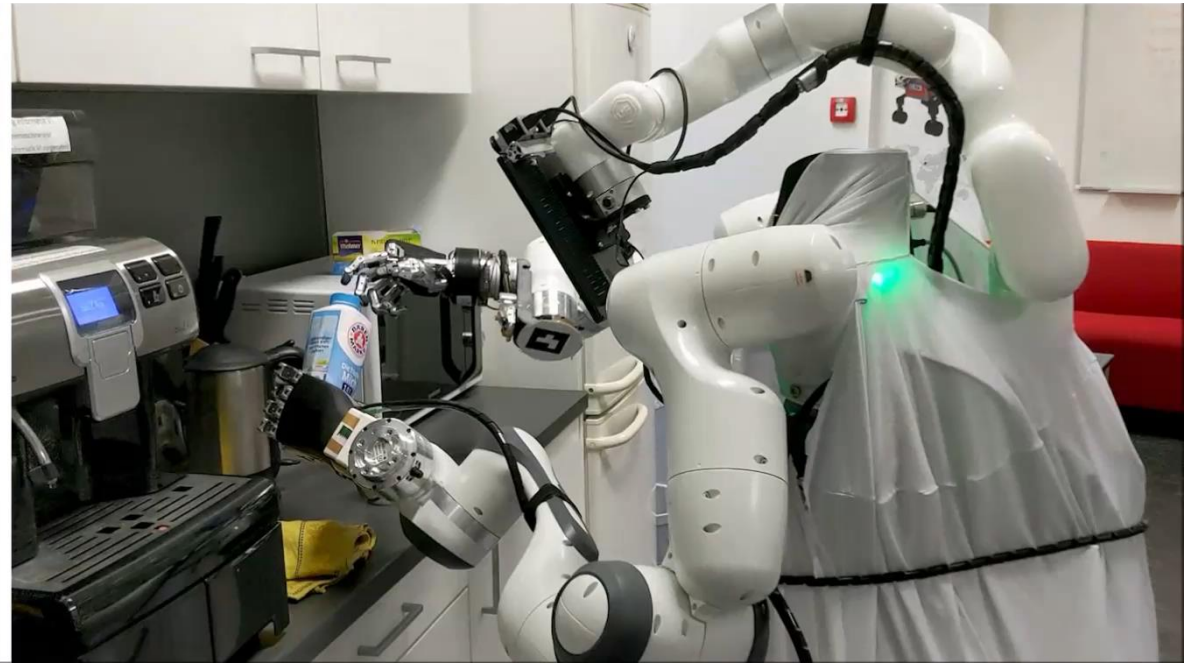
- Longer latencies require less direct control
  - Use autonomous skills, such as grasping an object or navigating to a waypoint
  - Shared autonomy where the operator controls high-level behavior and autonomy fills-in the low-level details (horse metaphor, Flemisch 2003)
- Operator might not always be available
  - 1:1 control often too costly
    ➔ one operator must supervise many robots
  - Issues of privacy and of being in operator‘s dept
- AI: Understanding intelligence by creating intelligent artefacts



[Photographer: Volker Lannert]

Unmatched Human Operators

- Humans can solve many tasks by teleoperation
  - Can cope with novel situations, quickly learn new tasks
  - Recognize and mitigate errors
- Far beyond the capabilities of autonomous robots



2x

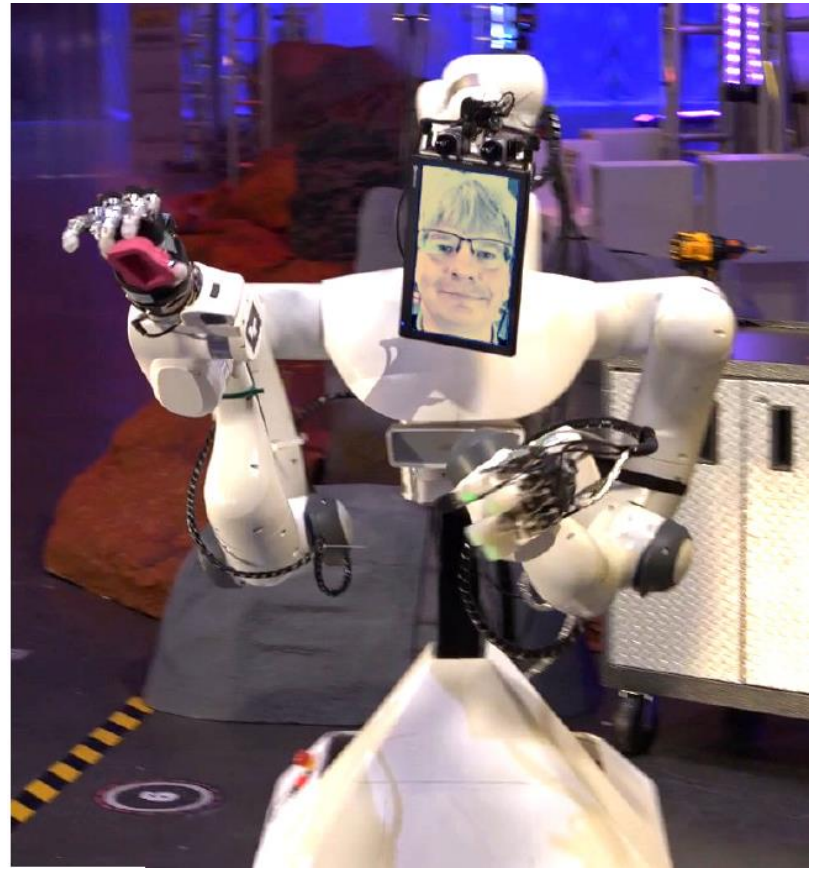
Team NimbRo Semifinal Submission

ANA AVATAR XPRIZE

UNIVERSITÄT BONN AIS

Conclusions

- The ANA Avatar XPRIZE competition advanced immersive telepresence systems
- Potential follow-up could raise the bar
  - Bandwidth restrictions and latencies
  - Locomotion on more difficult terrain
  - More complex manipulation (e.g., bimanual tasks)
  - Additional interaction modalities (e.g., temperature or smell)
  - Deeper interactions between avatars and recipients including interpretation of subtle communication cues and direct physical contact
- More autonomy is needed
- Need to match human cognitive functions



[XPRIZE]

Questions?