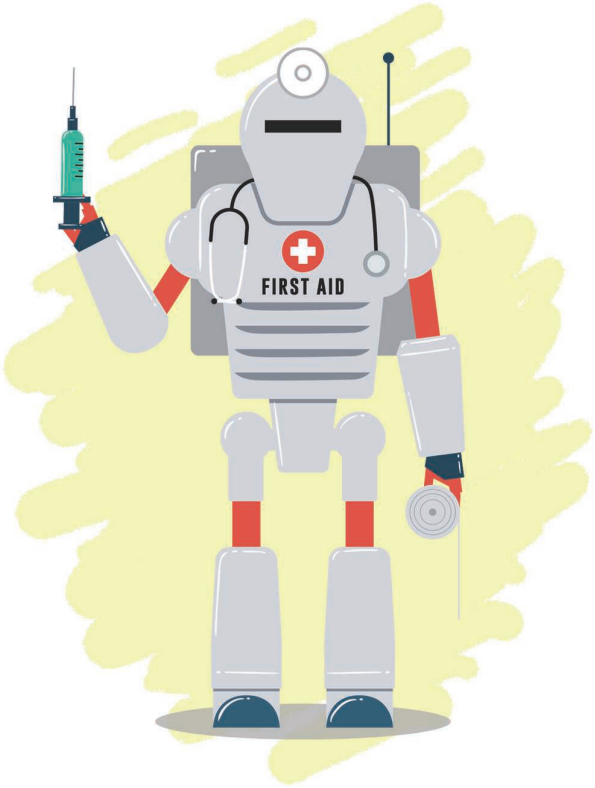


Flexible Disaster Response of Tomorrow

Final Presentation and Evaluation of the CENTAURO System

By Tobias Klamt, Diego Rodriguez, Lorenzo Baccelliere, Xi Chen, Domenico Chiaradia, Torben Cichon, Massimiliano Gabardi, Paolo Guria, Karl Holmquist, Malgorzata Kamedula, Hakan Karaoguz, Navvab Kashiri, Arturo Laurenzi, Christian Lenz, Daniele Leonardis, Enrico Mingo Hoffman, Luca Muratore, Dmytro Pavlichenko, Francesco Porcini, Zeyu Ren, Fabian Schilling, Max Schwarz, Massimiliano Solazzi, Michael Felsberg, Antonio Frisoli, Michael Gustmann, Patric Jensfelt, Klas Nordberg, Jürgen Roßmann, Uwe Süss, Nikos G. Tsagarakis, and Sven Behnke

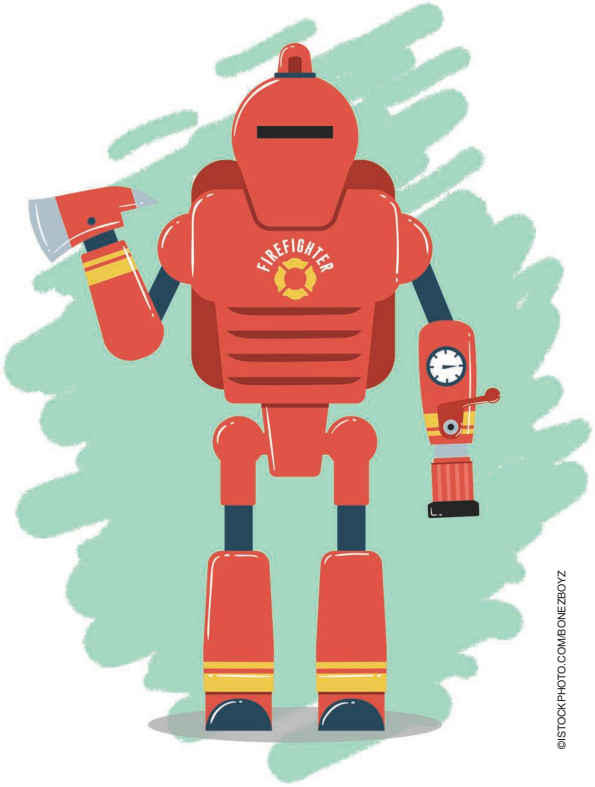
Digital Object Identifier 10.1109/MRA.2019.2941248
Date of current version: 21 October 2019

Mobile manipulation robots have great potential for roles in support of rescuers on disaster-response missions. Robots can operate in places too dangerous for humans and therefore can assist in accomplishing hazardous tasks while their human operators work at a safe distance. We developed a disaster-response system that consists of the highly flexible Centauro robot and suitable control interfaces, including an immersive telepresence suit and support-operator controls offering different levels of autonomy.

In this article, we describe the final CENTAURO system. We explain several high-level design decisions and how those were derived from requirements and the extensive experience of Kerntechnische Hilfsdienst, Karlsruhe, Germany (KHG), which is part of the German nuclear disaster-response organization and a potential end user of the system. We focus on components that were recently integrated and report about a systematic evaluation that demonstrated system capabilities and revealed valuable insights.

Disaster Response Needs Flexible Robots and Intuitive Teleoperation

Responses to natural and man-made disasters have attracted increased attention in the last decade because of their large potential impact. One example is the Fukushima disaster in March 2011. This was triggered when a tsunami hit the



Japanese coast—including the Fukushima Daiichi nuclear power plant—resulting in a nuclear disaster. There, rescue workers faced considerable risks to their health [1].

Although mobile manipulation robots have much potential to replace humans in such cases, the Fukushima nuclear disaster illustrated the gap between robot capabilities and real-world requirements. In such scenarios, a combination of locomotion and manipulation skills is required. Locomotion tasks include, for example, energy-efficient movement over long distances to approach a disaster site; navigation in man-made environments, such as along corridors, through doorways, onto ramps, and up staircases; and maneuvering over or around clutter, obstacles, and openings caused by the disaster. Possible manipulation tasks include moving heavy obstacles to clear paths, connecting and disconnecting plugs, unlocking and opening doors, and operating tools, such as power drills. Such robots also must incorporate their own energy resources, be operated out of sight of the operator, and be robust enough to function reliably over long periods in harsh conditions without maintenance.

We propose a set of operator interfaces with complementary strengths and different degrees of autonomy.

In recent years, advancements in the areas of actuation and sensing technology, computational hardware, and batteries have contributed to the development of robotic platforms that provide improved flexibility, precision, strength, robustness, energy efficiency, and compactness, enabling the deployment of robots in environments designed for humans. In the 2015 DARPA Robotics Challenge, teams were required to solve a set of mostly known locomotion and manipulation tasks. They demonstrated the state of the art for applied mobile manipulation robots. Successful robot entries included DRC-Hubo, Momaro, the CHIMP (Carnegie Mellon University Highly Intelligent Mobile Platform), and RoboSimian [2]–[5]. More recent mobile manipulation robots include RoboMantis [6] and Spot [7]. These possess flexible kinematic capabilities, but no methods to control these platforms have been presented. The humanoid Toyota T-HR3 [8] and the Walk-Man robot attached to a wheeled base [9] have been controlled through a head-mounted display (HMD) and by mirroring user movements to the robot. In the case of the T-HR3, additional force feedback is available. However, evaluation of these systems only addressed tasks of limited difficulty known in advance, so that operator training was possible.

The CENTAURO project aimed at creating mobile manipulation robots more useful for real-world missions. Hence, the development focused on a system that is flexible enough to solve a wide range of tasks considerably more challenging and realistic than those robots have performed before. Such tasks include those for which the operators had no chance to train the task, as in realistic disaster-response missions. We are convinced that a single operator interface that is flexible for use in a wide range of tasks—while being easy to use—does not exist. We propose a set of operator interfaces with complementary strengths and different degrees of autonomy. For tasks of limited complexity, high-level operator inputs suffice to control autonomous functionalities while relieving cognitive load from the operator. For rather difficult tasks, human experience and capabilities for scene understanding and assessing situations are indispensable. Thus, it is highly desirable to transfer these cognitive capabilities into the disaster scene through such technologies as, for example, a telepresence suit.

Many technical details of the system have been described in [10], a report on the intermediate system status one year before the end of the project. In this article, we present the final system at the end of the project and explain why we made several fundamental design decisions. Furthermore, we describe components that have been recently developed and, thus, are not included in the report, such as a new robust end effector, a new motion controller, a bimanual control interface for the telepresence suit, technology for bimanual autonomous grasping, a terrain classification method for autonomous locomotion, and virtual models for the robot and its environments. Finally, we present a systematic evaluation that goes significantly beyond previously reported experiments. It consists of tasks that are considerably more challenging so that the system can be fairly assessed for potential real-world applications.

The Centauro Robot

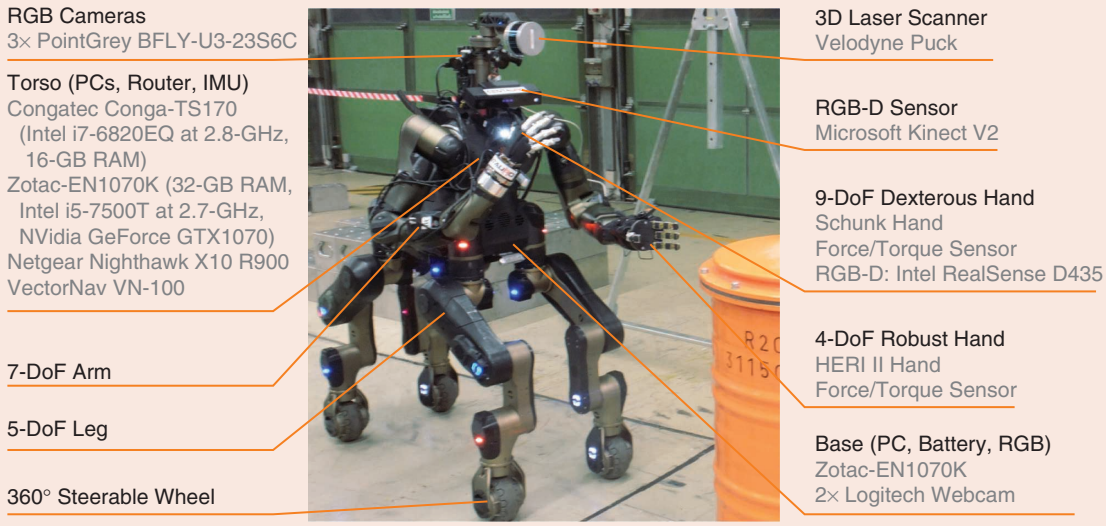
In designing the robot, we looked at requirements of typical disaster-response missions and experiences of KHG, which operates a fleet of disaster-response robots. Those platforms are mostly tracked vehicles with a single arm ending in a 1-degree-of-freedom (DoF) gripper. Moreover, additional monochromatic or red, green, blue (RGB) cameras enable the operator to view the scene.

Hardware Description

Regarding the locomotion concept, KHG reported that tracked approaches are robust and easy to control but exhibit difficulties in very challenging environments, such as in places where large debris block a path or where there are large gaps or holes to traverse or avoid. Moreover, the lower energy efficiency of tracks means that the systems must have large batteries, thus increasing weight. Wheeled platforms generally offer higher energy efficiency, but their mobility is limited to sufficiently flat terrains. Meanwhile, legged robots show promise of being able to traverse challenging terrains (including steps and gaps) since only isolated footholds are needed. For the Centauro robot (Figure 1), we chose a hybrid driving–stepping design using legs ending in wheels. In this manner,

the complementary strengths of the two locomotion approaches are combined. Furthermore, the robot can switch between locomotion modes without changing its posture, which enables additional unique motions, such as individual feet movement relative to the base while keeping ground contact—a valuable property for balancing. Wheels are 360° steerable and actively driven, allowing for omnidirectional driving. The legs incorporate 5 DoF each, with three upper joints arranged in a yaw–pitch–pitch configuration and two lower joints for wheel orientation and steering. When stepping, wheel rotation is blocked so that the grip is similar to point-like contacts of legged robots. We chose a design with four legs due to its higher static stability (support polygon area), compared to bipedal robots.

KHG also reported that bimanual setups can be advantageous in several situations but are rarely present on the market. Furthermore, operators sometimes experience problems understanding kinematic constraints of the manipulator arms. Motivated by these observations, we chose an anthropomorphic upper-body design for Centauro, resulting in an eponymous centaur-like body plan. It enables effective bimanual manipulation in human workspaces. A yaw joint connects the upper body to the lower body. The two 7-DoF robot arms possess a quasi-anthropomorphic profile, providing an intuitive kinematic system understanding for the operators. To obtain versatile grasping capabilities, two different end effectors with complementary strengths are incorporated. On the right, a commercial, anthropomorphic, 9-DoF Schunk SVH hand provides dexterous grasping



(a)

Robot Specifications	
Height	112–171 cm
Width	61 cm (min)
Length	61 cm (min)
Weight	92 kg (Including Battery)
Number of DoF	55 (Including Hands)
Wheel Diameter	15.6 cm
v_{max} Driving	1.6 m/s
Battery Capacity	34.6 Ah
Battery Voltage	48 V
Battery Current (Maximum)	80 A
Battery Weight	7.5 kg
Operation Time	~2 h

(b)

Figure 1. The Centauro robot. (a) Centauro approaching a cask to conduct a radiation measurement. Labels describe the hardware components and their location. (b) Robot hardware specifications. IMU: inertial measurement unit.

capabilities. On the left, a flexible, 4-DoF Hardware Embedded Reduced Intricacy (HERI) II hand [11] allows for a higher payload and exhibits higher compliance and robustness. The HERI II hand was developed specifically to overcome the limitation of typical underactuated hands in executing basic dexterous motions, such as pinching and triggering tools, while maintaining the main advantages of underactuated designs.

With respect to the sensor setup, KHG reported that the operation of its robots, based solely on camera images, puts a high cognitive load on the operators and requires extensive training to derive correct situation assessments from those images. The lack of assessment of the overall kinematic system state and of the robot's environment causes uncertainty and stress. Therefore, Centauro incorporates a multimodel sensor setup that provides comprehensive visualizations to the operators and input to (semi)autonomous control functionalities. It includes a 3D laser scanner with spherical field of view, an RGB depth (RGB-D) sensor, and three RGB cameras in a panoramic configuration at the robot head. A pan-tilt mechanism allows for adaptation of the RGB-D sensor pose, facilitating manipulation at different locations. Furthermore, an inertial measurement unit (IMU) is mounted in the torso. Two RGB cameras under the robot base provide views on the feet. An RGB-D sensor at the right wrist creates an additional perspective during manipulation. Finally, two 6-DoF force/torque sensors are mounted between the robot arms and hands.

A semiautonomous stepping controller is used for walking across irregular terrain.

To the best of our knowledge, no system on the market provides hybrid locomotion combined with bimanual manipulation and the required sensor setup. The robot has been mainly fabricated of an aluminum alloy. To drive the robot, five sizes of torque-controlled actuators have been developed. The compactly integrated series elasticities provide mechanical robustness to impacts and torque sensing via deformation monitoring. Based on an effort analysis, a suitable actuator was chosen for each joint [12]. In addition, the robot body houses a Wi-Fi router, a battery, and three PCs for real-time control, perception, and higher-level behavior.

One consideration in designing a disaster-response robot is to include hardening to withstand radiation. This can be realized through either massive lead cases for the computational hardware, which would add significant weight, or the use of radiation-immune computer hardware, which is severely limited, compared to commercial off-the-shelf (COTS) hardware. We developed the CENTAURO system as a scientific demonstrator that aims at pushing the state of the art. Hardening the system against radiation would have affected our research by considerably limiting the modular hardware design and the use of cheaper COTS components. We thus see system hardening as a task for further development according to the demands of the end user.

Software Description

For low-level hardware communication, we developed XBotCore, an open-source, real-time, cross-robot software platform [13]. This software makes it possible to seamlessly program and control any robotic platform while providing a canonical abstraction that hides the specifics of the underlying hardware. Higher-level motion commands are processed in the developed CartesI/O motion controller [14], which can consider multiple tasks and constraints with given, situation-specific priorities. The controller resolves a cascade of quadratic programming (QP) problems, each associated with a priority level, and considers the optimality reached in all of the previous priority levels. The high computational cost of the successive QP problems is alleviated by assessing our inverse kinematics (IK) loop in a range between 1,000 and 250 Hz.

The Operator Station

KHG reported that during its real-world missions, human concentration suffers from a high cognitive load. Thus, operators are usually exchanged approximately every 15 min. The kinematics of Centauro are considerably more complex, which means the cognitive load would likely be higher. In discussions with KHG, we identified several reasons for these high cognitive loads.

1) Camera images are not comprehensive enough to enable operators to fully understand the robot's situation, especially if contact with the environment has to be assessed.
2) Operators have problems understanding the robot kinematics if they cannot see them and cannot assess joint limits and self-collisions.
3) Operation of the robot requires continuous direct control of all robot capabilities, which demands an exhausting, continuous, and high level of concentration.

We derived several requirements for the CENTAURO operator station from these experiences. Regarding visualization, we aimed at providing a more intuitive understanding of the situation through a 3D virtual model of the robot and its environment. This model is generated from joint states, the IMU, the laser scanner, and the RGB-D sensors. It is enriched with camera images from several perspectives. We refer to these digital counterparts of the robot and its environment as *digital twins*. Moreover, we added force feedback to improve situational assessment, especially if environment contacts are involved. Regarding control, several interfaces with different strengths were implemented. The main operator intuitively controls the robot through a full-body telepresence suit, while support operators are able to command additional control functionalities on different levels of autonomy.

Full-Body Telepresence Suit

The main part of the full-body telepresence suit is the upper-body exoskeleton with arm exoskeletons covering shoulders and elbows, two spherical wrist exoskeletons, and two hand exoskeletons (Figure 2). Hence, complex operator arm, wrist, and hand motions can be intuitively transferred to the robot. While the arm motions can be also realized through off-the-shelf lightweight manipulators, the exoskeleton, designed to maximize performance in human–robot interaction, features high transparency, a large workspace, and high maximum payload/rendered force. In contrast to similar teleoperation systems [15], the presented exoskeleton considers force feedback of both the arm and the hand–palm. The arm exoskeletons cover about 90% of the natural workspace of the human upper arms without singularities. To keep the number of moving masses low, the device is driven by electric actuators remotely located behind the operator seat and connected to the joints through idle pulleys and in-tension metallic tendons. Force feedback is provided for each individual joint. Two rotational 3-DoF forearm–wrist exoskeletons track operator wrist motions and provide independent torque feedback for each axis [16]. In developing these components, researchers focused on making the exoskeleton comfortable for the person wearing it and maintaining an open structure to avoid collisions between parts during bimanual operational tasks. The 6D end-effector poses of the two arms are computed and sent to the Centauro robot, which uses them as references to compute joint values through IK. Two hand exoskeleton modules track operator finger motions and provide grasping-force feedback [17]. Figure 3 shows the overall architecture of the telepresence suit. The control incorporates friction and gravity compensation as well as the time-domain passivity approach combined with a position-force scheme to achieve high transparency [18].

Since Centauro's quadrupedal lower body is considerably different from the bipedal lower body of humans, direct control of all four legs through an exoskeleton is not feasible. Instead, we decided to use pedals to control omnidirectional driving, while more challenging locomotion is controlled by the support operators. In addition, the operator wears an HMD for immersive 3D visualizations.

Support Operator Interfaces

While the telepresence suit is especially useful for tasks requiring challenging manipulation, several other tasks can be managed without the suit and with an increasing level of autonomy—requiring less input from the operator. Omnidirectional driving can be controlled through a joystick. A 6D mouse provides Cartesian control of 6D end-effector poses for arms and legs. For more complex motions, a keyframe editor enables configuration control and motion generation in joint space and in Cartesian end-effector space.

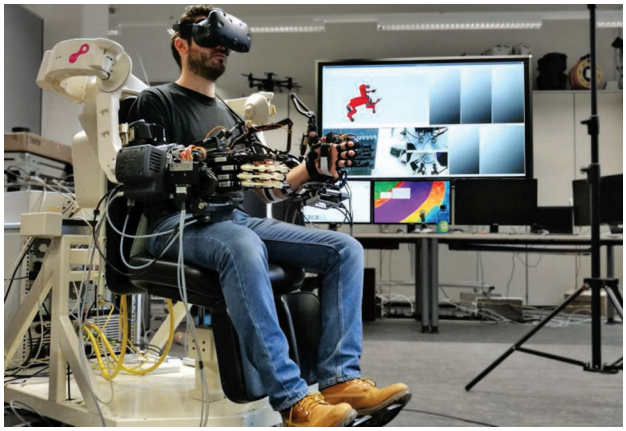


Figure 2. The telepresence suit allows the operator to intuitively control the remote Centauro robot in bimanual manipulation tasks while providing force feedback. The support operator station can be seen in the background.

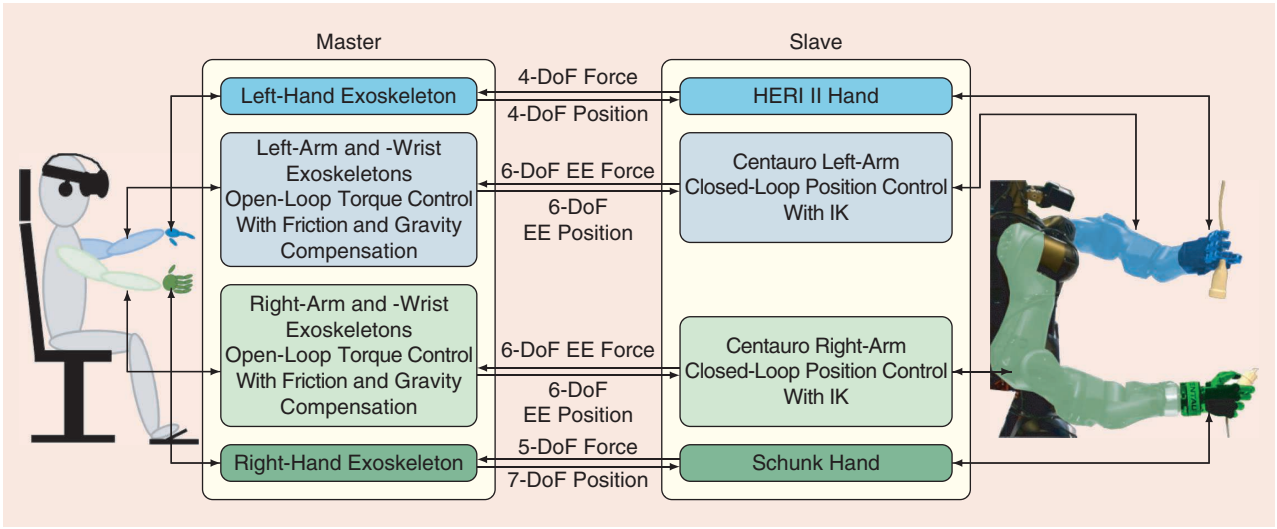


Figure 3. An illustration showing the multilateral position–force teleoperation architecture. While hand positions are perceived on joint level and hand forces are measured for each finger, arm positions and forces are projected to the end effectors (EEs).

A semiautonomous stepping controller is used for walking across irregular terrain. The operator can trigger individual stepping motions, while the controller automatically generates a statically stable robot configuration and a subsequent foot-stepping motion sequence, including ground contact estimation. The operator can also define goal poses for an autonomous driving–stepping locomotion planner (see the section "Autonomous Hybrid Driving–Stepping Locomotion"). Bimanual manipulation can be controlled through the 6D mouse interface, through the keyframe editor, or by defining targets for autonomous grasping (see the section "Autonomous Dual-Arm Manipulation").

Situation Prediction With Digital Twins

A digital twin virtually represents a real object, including its physical and kinematic properties, appearance, data sources, internal functions, and communication interfaces. We use the Verosim simulation environment for the digital twin (in the loop with the real robot).

In addition, the robot can be placed in simulated environments to support development. Virtual sensor models can generate detailed measurements of the virtual environment, and physical interactions of rigid bodies are simulated. The digital twin interfaces to external control and processing units are identical to those of the real robot [19].

The robot's digital twin is accompanied by a digital twin of the robot environment. Three-dimensional meshes are generated from laser scanner measurements and can be overlaid by texture coming from, for example, RGB cameras. A large model database contains many static objects, such as steps or plugs, and dynamic objects, such as valves or doors. If such objects are recognized, virtual models of these objects can automatically be positioned in the virtual environment. Hence, limitations of sensory input, which are caused by occlusion, for example, are mitigated, resulting in a holistic, detailed virtual representation of the scene. This representation is used to generate immersive user visualizations for both 3D HMDs and the 2D renderings from arbitrary view poses [20]. Visualizations can be enriched with control-interface-specific data, such as locomotion paths, which can be inspected before execution.

During robot operation, the robot might become involved in a situation so challenging that the operator is uncertain how the robot should proceed. Since the digital twins of the robot and its environment accompany the mission and are frequently updated, the same situation is represented in the virtual environment. A predictive mode is activated by freezing the real robot in its current configuration and decoupling the digital twins from their real counterparts. Subsequently, the operator control is redirected from the real robot to its digital twin (Figure 4). Consequently, operators interact with the virtual representation as they would with the real robot. This can be used to, for example, evaluate

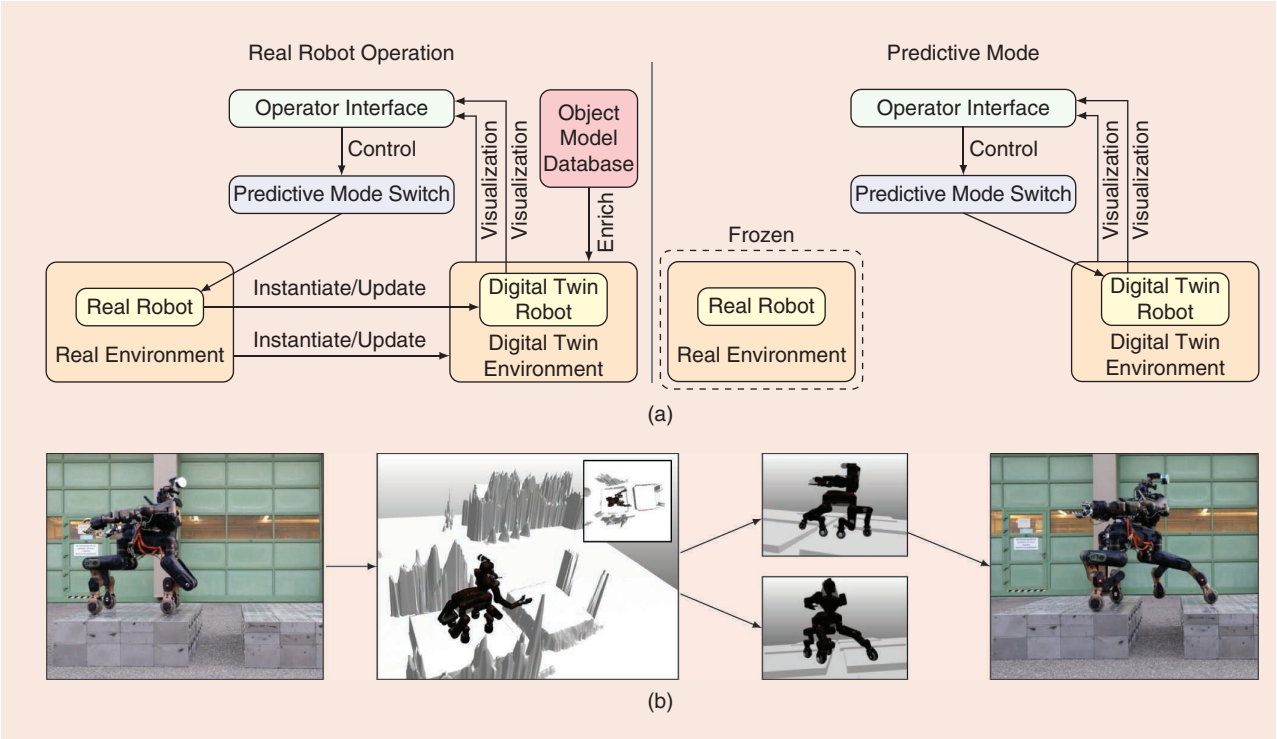


Figure 4. Diagrams, photos, and illustrations showing the use of a digital twin in predictive mode. (a) The communication architecture for real robot operation and the predictive mode. (b) An illustration and photos showing how, when Centauro arrives at the gap, the digital twins of the robot and its environment represent the same situation. The operators switch to the predictive mode and try out different strategies to traverse the gap while the real robot remains in place. Eventually, the real robot can execute the identified solution.

multiple approaches to solve a specific task. Once a satisfying solution is found, the task can be executed with the real robot [21].

Operator Relief Through Autonomous Control Functions

While intuitive teleoperation is superior for solving complex, unknown tasks, disaster-response missions also include several tasks that can be automated to reduce the high cognitive load on the operators. Such tasks might include navigating to a desired location or grasping a tool. Moreover, in comparison to direct teleoperation, autonomy is often faster and depends less on a reliable data link. However, the development of autonomous skills is challenging due to the level of variability of the environment and the tasks.

Autonomous Hybrid Driving–Stepping Locomotion

The many DoF that have to be controlled in combination with the required precision and balance assessment put a high cognitive load on the operator and might result in motion too slow to be practical. We developed a locomotion planner that combines omnidirectional driving and stepping capabilities in a single graph-search-based planning problem [22]. The environment is represented with 2D height maps computed from laser scanner measurements. These maps are processed to estimate costs for the robot base and individual feet, enabling precise planning.

Two-dimensional height maps are inadequate for representing terrains, such as expanses of fine gravel. Therefore, the planning pipeline is enriched with terrain classification providing additional semantics (Figure 5). A geometry-based analysis derives the terrain slope and roughness from point clouds. In parallel, a vision-based terrain classification is performed on RGB images by employing a convolutional encoder–decoder neural network. Initial training on the CityScapes data set [23] provides the network with reliable representations of drivable surfaces, walls, vegetation, and terrain. Subsequent training on custom data sets from forests and buildings refines the terrain representations. We focused on the classification of staircases, which have many similarities to obstacles but differ in traversability. Finally, all features are fused to pixelwise traversability assessments (safe/risky/obstacle/stair), which are incorporated in the planner as an additional cost term.

The fine planning resolution (2.5 cm) and high-dimensional (7 DoF) robot representation result in rapidly growing state spaces that exceed feasible computation times and available memory for larger planning problems. We improved the planner to generate an additional coarse, low-dimensional (3 DoF), and semantically enriched abstract planning representation [24]. The cost function of this representation is learned by a convolutional neural network trained on artificial short planning tasks but generalizes well to real-world scenes. An informed heuristic (containing knowledge about the environment and the robot) employs this abstract representation and exploits it to effectively guide the planner toward the goal. This accelerates planning by multiple orders of magnitude with comparable result quality.

Autonomous Dual-Arm Manipulation

While object grasping could be executed through the telepresence suit, autonomous grasping is promising to relieve

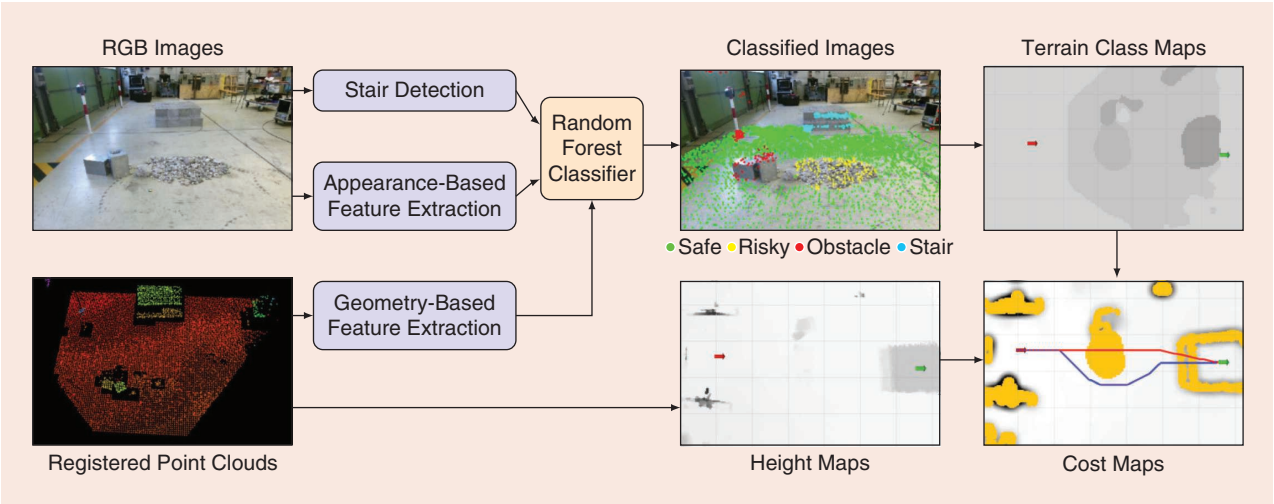


Figure 5. Photos and a chart showing how an autonomous locomotion cost map is generated. Point clouds are processed to height maps to finally generate cost maps. This pipeline is enriched by parallel terrain class computation. Geometry- and appearance-based features are extracted from point clouds and RGB images. Stairs are detected in an additional module. Outputs are merged in a random forest classifier. Terrain classes are added to the cost computation. Note that the patch of gravel in front of the robot (red arrow) is not recognized well in the height map. The terrain class map clearly represents this area and enriches the cost map with this information. The cost map also shows two generated paths to a goal pose on top of the staircase (green arrow). The operator can choose one of these paths. The red path represents a low weight for terrain class-based costs: the robot traverses the patch of gravel. The blue path incorporates terrain class-based costs with a higher weight: the robot avoids the gravel.

cognitive load from the operator. Some objects, such as tools, require functional grasps: they must be grasped in a specific way to enable their use. This requirement poses challenges to the autonomy. Additional challenges arise for bimanual tasks, which force the execution of collision-free synchronous grasping motions.

The presented autonomous grasping pipeline enables bimanual functional grasping of unknown objects if other instances of this object class are known (Figure 6). Input is RGB-D data. First, semantic segmentation—based on Refine-Net with additional postprocessing—finds object contours and outputs corresponding point clouds. It is trained on synthetic scenes with objects coming from turntable captures or CAD mesh renderings. Subsequently, the object pose is estimated from the extracted point cloud, and a grasp is generated. Grasp generation is based on a nonrigid registration that incorporates class-specific information embedded in a low-dimensional shape space for the object category. To learn the shape space of a category, object point clouds are registered toward a single canonical instance of the category [25], [26]. During inference, we optimize a shape that best matches the observed point cloud. Finally, individual grasping motions for each arm, which are known for the canonical model, are warped to fit the inferred object shape. Arm trajectories are subsequently optimized by incorporating collision avoidance, joint limits, end-effector orientation, joint torques, and trajectory duration [27]. Since the optimization is performed in joint space, multiple end effectors can be easily added to the optimization problem by incorporating additional terms in the cost function. For manipulating an object (once it is grasped) with multiple end effectors simultaneously, an additional term that penalizes violations of the closed kinematics chain constraint is added. The planned motion is finally presented to the operator for verification. The execution of the complete planning pipeline takes fewer than 7 s.

Two-dimensional height maps are inadequate for representing terrains, such as expanses of fine gravel.

Evaluation

The final CENTAURO system was evaluated using a systematic benchmark test at the facilities of KHG. Tasks were designed based on KHG's knowledge about real-world disaster-response missions. While the intermediate evaluation reported in [10] was carried out at the same facilities, one year of further intensive system development significantly extended and advanced the CENTAURO system. Hence, the described tasks in this article are designed to be considerably more challenging to demonstrate the significantly increased real-world applicability of the final CENTAURO system.

All tasks were performed without previous training and with the operator station located in a separated room, thus preventing direct visual contact. The robot operated mostly untethered, relying on power supplied through its battery and

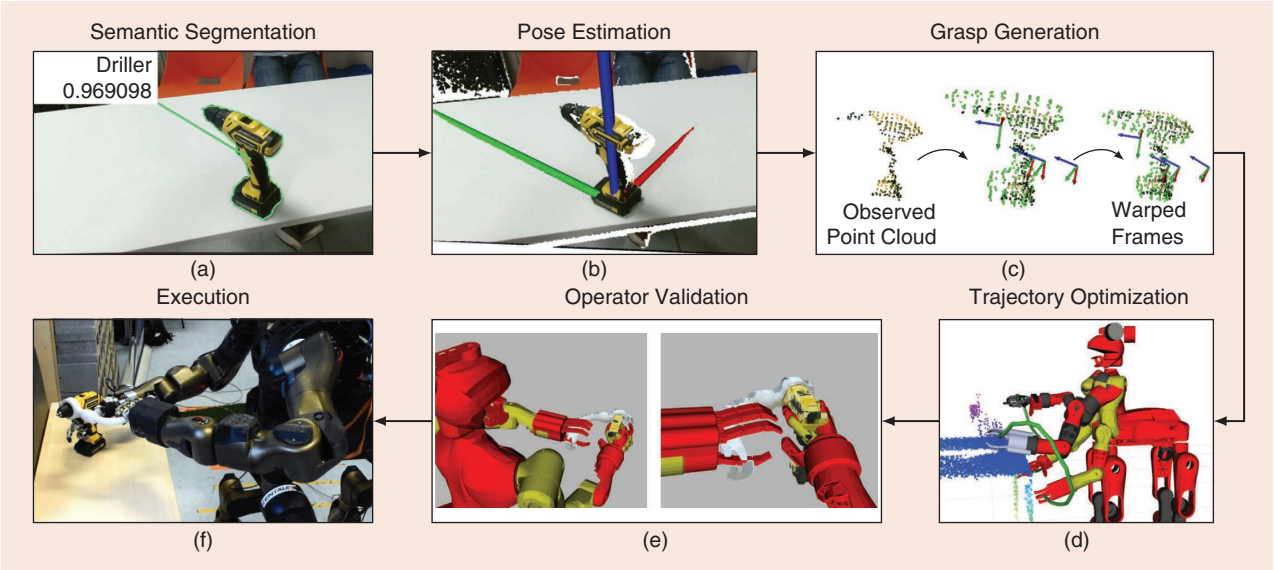


Figure 6. Photos and illustrations showing the autonomous manipulation pipeline. (a) From RGB-D images perceived by the robot, the semantic segmentation finds the contour of the object—a drill, in this case—and (b) extracts the point cloud. Based on the contour and the point cloud, the 6D pose of the object is estimated. (c) Subsequently, the segmented point cloud is registered nonrigidly using the shape space of the object category, and control poses (e.g., pregrasp and grasp poses) for both arms are warped from the grasping motion of the canonical object. (d) The final joint trajectory is generated by a trajectory optimizer that guarantees collision-free motions. (e) The resulting motion is verified by the operator and (f) finally executed.

a wireless data link. A neutral referee rated how well the task was performed and ensured that no communication took place between persons at the robot location and the operators. Since the system had research demonstrator status and this evaluation was the first time all required components were integrated into a disaster-response system, we permitted up to three runs for each task, of which only the best run was rated. Nevertheless, most tasks were conducted successfully at the first trial without further runs.

The tasks are not described in chronological order but in categories. During the evaluation, the Platform+Lever task was one of the first manipulation tasks to be performed. Unfortunately, a mechanical part in the upper-body exoskeleton broke during this task, and we were not able to repair it in the limited time of the evaluation meeting. Therefore, we opted to solve all remaining tasks with the remaining interfaces. However, in the section "Isolated Evaluation of Telepresence Suit Control," we present an isolated evaluation of the telepresence suit.

The task-specific performance of the evaluation at KHG is summarized in Table 1. Impressions of the experiments are given in Figure 7. Further video footage is available online (https://www.ais.uni-bonn.de/videos/RAM_2019_Centauro).

Locomotion Tasks

Locomotion capabilities were evaluated in three tasks. In the Step Field task, the robot had to traverse a surface of concrete blocks in a random grid scheme. The task difficulty was considerably increased by wooden bars randomly placed on the surface. We used the semiautonomous stepping controller and the joystick for omnidirectional driving to successfully solve this task. Manual adaptations through the 6D mouse and the keyframe editor were frequently required, due to the irregularities caused by the wooden bars. In addition, the RGB camera under the robot base lost connection. It normally provides a detailed view on the rear feet, which is essential for situational assessment. To compensate, the right-arm RGB-D sensor was used when executing individual rear-foot motions.

In the Staircase task, the robot had to climb a staircase with three steps. We used predefined motion primitives and joystick driving control to accomplish this task. Once the robot was on the staircase, motions could be triggered in a repetitive manner. Hence, climbing longer staircases would only require more repetitions of these motions.

In the Gap task, the robot had to traverse a 50-cm gap. We used motion primitives and joystick commands to

Table 1. Results of the KHG evaluation.

Task	Rating (%)	Time (min:s)
Locomotion tasks		
Step Field: navigating an uneven surface strewn with random debris	100	87:38
Staircase: climbing a flight of stairs	100	10:02
Gap: overcoming a 50-cm gap	100	3:26
Manipulation tasks		
Fire Hose: connecting and disconnecting a fire hose and nozzle using a special wrench	100	10:10
230-V Connector (Standard): plugging a standard 230-V plug into a cable outlet	100	6:50
230-V Connector [Certification of Electrotechnical Equipment (CEE)]: plugging a CEE-type plug into an outlet with a lid	100	10:00
Shackle: fixing a shackle to a metal ring at the wall by inserting and turning a screw	100	24:35
Electric Screwdriver: driving a screw with an electric screwdriver into a wooden board (held with one hand) to attach the board to a wooden block	100	6:21
Power Drill: using a two-handed power drill to make holes at marked places in a wooden wall	90	2:50
Combined tasks		
Door: unlocking and opening a conventional door with a handle and passing through the doorway	100	13:30
Platform+Gate-/Lever-Type Valve: approaching and climbing a platform with the front feet and opening and closing a gate-/lever-type valve	60/100	23:30/6:50
Grasp/Visualize Pipe Star: grasping/visualizing defined positions of a pipe star on the ground	95/100	14:25/6:09
Autonomous tasks		
Autonomous Locomotion: autonomously navigating around debris and climbing stairs to reach a goal pose on an elevated platform	75	2:50
Autonomous Manipulation: autonomously grasping an unknown, two-handed power drill	75	0:46

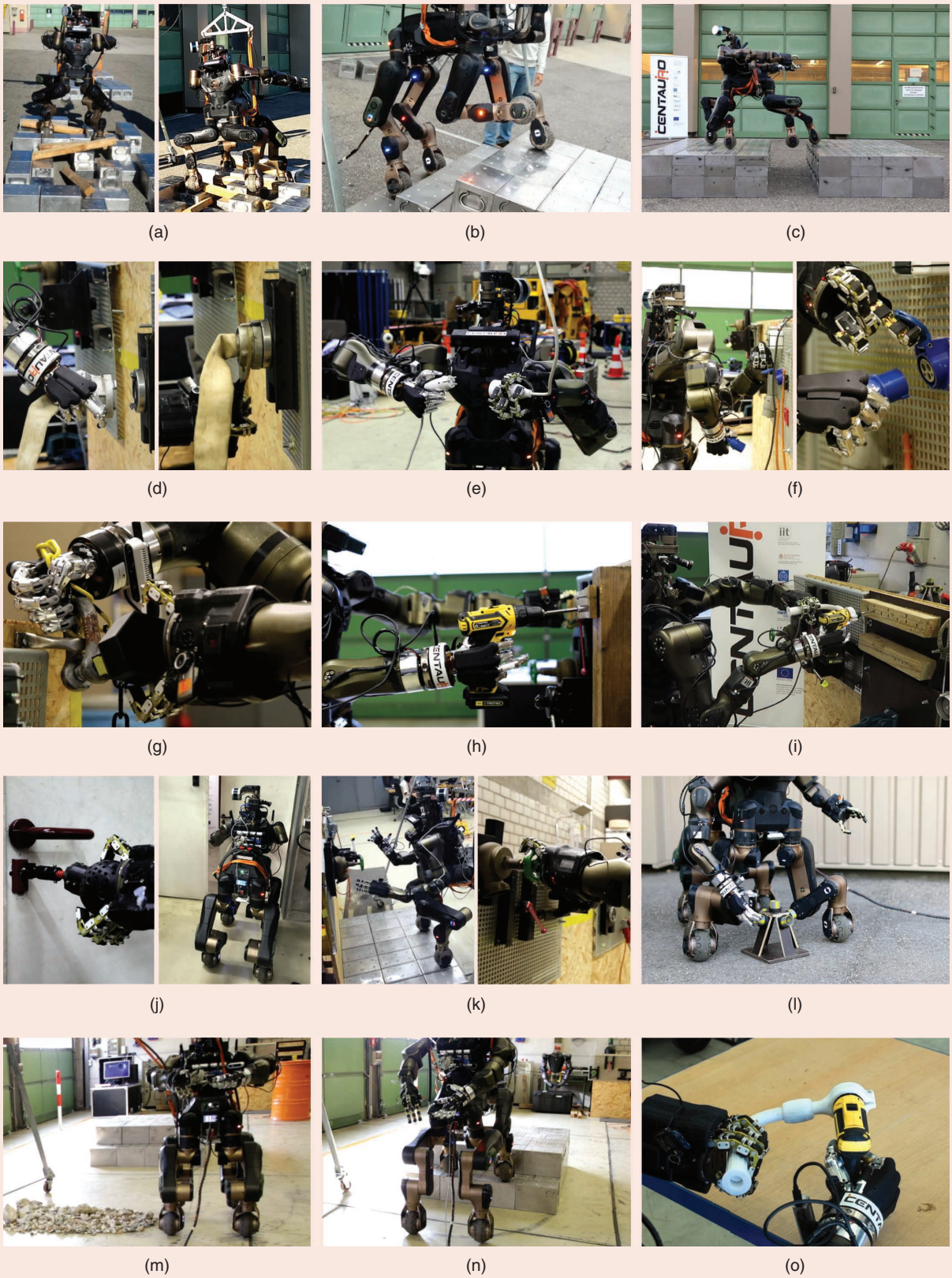


Figure 7. Impressions of the final evaluation at KHG. The locomotion tasks: (a) Step Field, (b) Staircase, and (c) Gap. The manipulation tasks: (d) Fire Hose, (e) 230-V Connector (Standard), (f) 230-V Connector (CEE), (g) Shackle, (h) Electric Screwdriver, and (i) Power Drill. The combined tasks: (j) Door, (k) Platform and Valve, and (l) Pipe Star. The autonomous tasks: (m) and (n) Hybrid Driving–Stepping Locomotion and (o) Bimanual Manipulation.

successfully accomplish this task. In addition, the predictive mode of the digital twin was employed to evaluate different strategies (Figure 4). That mode was demonstrated separately and is not included in the measured time.

Manipulation Tasks

Disaster-response manipulation capabilities were evaluated in six tasks. In the Fire Hose task, the robot had to connect and disconnect a fire hose with a nozzle. Integrated sealing caused high friction, requiring the use of a designated wrench. Dexterous grasping capabilities of the Schunk hand were required to mount the fire hose on the nozzle. The HERI II hand operated the wrench. This task was accomplished using the 6D mouse control. Its feature enabling only certain motion axes facilitated precise alignment.

In the 230-V Connector (Standard) task, the robot had to connect and disconnect a standard 230-V plug placed in one hand to a cable power outlet hanging from the ceiling. After grasping the power outlet, precise alignment was challenging since vision was hindered through the enclosure of the power outlet. In a second task version, the robot had to plug a 230-V CEE-type plug, which was closed with a lid incorporating a spring mechanism, into a power outlet at the wall. One hand was required to dexterously open the lid and keep it open while the other hand had to insert the plug. This task was challenging due to potential self-collisions between hands as well as several occlusions. For both tasks, the 6D mouse control was well suited since it allowed very precise motions. However, force feedback was not present, and contact had to be assessed from visualizations.

In the Shackle task, the robot had to fix a shackle to a metal ring at the wall. One hand had to position the shackle around the ring, while the other hand had to insert and turn a screw. An adapter for the small screw head facilitated grasping and turning. The 6D mouse interface was used, and its precise motions were required to align the screw with the thread.

In the Electric Screwdriver task, the robot had to screw a wooden board to a wooden block using an off-the-shelf electric screwdriver. The wooden board with premounted screws was held with the HERI II hand. The anthropomorphic Schunk hand operated the electric screwdriver using its index finger for triggering. The 6D mouse control was used to align the tip of the screwdriver with the screw head. When activating the screwdriver, the screw moved inside the wood, requiring fast reaction to follow the screw head toward the wooden board.

Finally, the Power Drill task demonstrated the operation of a bimanual tool by making holes at marked places into a wooden block. Again, the Schunk hand was used to pull the trigger of the drill. The 6D mouse control was set to generate bimanual motions with respect to a custom frame in the tool. Since one hole was drilled with a deviation of 1 cm to the marked position, this task achieved a rating of 90%.

Combined Tasks

Three tasks combined locomotion and manipulation. They were designed to represent realistic missions and to evaluate the coordination of different control interfaces.

In the Door task, the robot had to unlock and open a locked door and pass through the doorway. The key was mounted on an adapter that KHG uses in real-world missions. Once the key was inserted in the lock, multiple regrasps were necessary to turn it. This was achieved by 6D mouse control. After opening the door, the arms were moved to a configuration so that they further opened the door when driving through it. Omnidirectional driving was controlled by joystick.

In the Platform+Valve task, the robot had to approach an elevated platform, climb on it with its front feet, and open and close a valve. This task was performed with gate- and lever-type valves. Approaching the platform was realized by joystick, while the semiautonomous stepping controller was employed to climb the platform. Centauro's upper body was controlled through the telepresence suit. The gate-type valve was turned 360° in each direction for opening and closing. While the operators thought that this rotation was sufficient, the referee expected full opening and closing (>360°). A rating of only 60% was given. Subsequently, when operating the lever-type valve, a mechanical part broke inside the upper-body exoskeleton after closing the valve 30°. To demonstrate the robustness of our system in case of failure, we switched to 6D mouse control on the fly and finished the task.

To demonstrate the robustness of our system in case of failure, we switched to 6D mouse control on the fly and finished the task.

Finally, the Pipe Star task was performed. A pipe star was placed on the ground. It consists of five short pipes with different orientations and different symbols inside. Each of the pipes had to be grasped at the top, and the symbol inside the pipe had to be perceived. This task is used by KHG to evaluate real-world mobile manipulation platforms. Initially, the keyframe editor was used to move the robot to a very low base pose, allowing for manipulation close to the ground [Figure 7 (l)]. Operators used the joystick for omnidirectional driving and the 6D mouse for arm motions. When grasping the pipes, the top pipe was touched from the side only and not from the top. While operators thought they had accomplished the task, the referee rated this only as a partial success since he determined that the task required contact from the top. Moreover, it was challenging for the operators to assess physical contact since force feedback was not available. Visualizing the symbols inside the pipes was realized through the wrist RGB-D camera without problems.

Autonomous Tasks

Autonomous control functionalities were evaluated in two tasks. In the Autonomous Locomotion task, the robot started in front of a patch of debris and some obstacles. A two-step staircase ending in an elevated platform was positioned behind the debris with the goal pose on the platform (Figure 5). Map generation, localization, goal-pose definition, and planning took around 30 s. The robot detected the debris, drove around it, and arrived at the staircase. It then started to climb the staircase. However, after climbing two steps with its front feet and then attempting to climb the first step with a rear foot, the robot lost its balance. Essentially, the robot model, which was used by the planner, did not include recent hardware modifications (such as integration of the battery) and, hence, based its balance assessment on a wrong center of mass.

In the Autonomous Manipulation task, the robot had to detect an unknown bimanual power drill on a table in front of it, grasp it with both hands, operate the trigger, lift it, and put it back on the table. The robot successfully detected the drill and derived a valid grasping pose, which it approached while avoiding collisions with the table. However, due to a small misalignment of the right-hand grasp, it failed to activate the tool and also touched the tool after placing it back on the table, causing the tool to fall.

Isolated Evaluation of Telepresence Suit Control

The telepresence suit control was evaluated in isolated experiments later. Again, direct visual contact between the operator and the robot was prevented. In a first task, a wrench had to be grasped, and a valve stem had to be turned. This task included fine positioning and orientation control of the tool tip and compliance and modulation of exerted forces on specific axes. Force feedback enabled the operator to feel the interaction between the wrench and the valve stem, facilitating positioning of the tool.

In a second task, bimanual coordination was evaluated. The operator had to grasp a glue gun positioned in an unfavorable position on a table in front of the robot. Since a single direct grasp would not allow for correct tool usage, the operator had the robot pick up the tool with the HERI II hand, hand it over to the Schunk hand, and adjust it to obtain a functional grasp that made it possible to press the trigger. This task required precise position and orientation control of both hands, combined with adequate grasp forces to avoid slippage. In addition, problems derived from the force loop were alleviated by the compliance of the controller, creating stable behavior without noticeable oscillations on either the operator or the robot side. The task took 88 s to complete.

The telepresence suit control was fast enough to reliably solve a challenging task in a reasonable time, but it was considerably slower than a human executing the task. Two reasons were identified:

1) The operator received no feedback for collisions between the robot elbow and lower body. Introducing another modality, such as auditory feedback, could help the operator.
2) The robot hands often caused considerable occlusions in the 3D operator visualization when manipulating objects. We observed that the main operator relied only on 3D visualizations if those were available in high quality and with few occlusions. Otherwise, RGB camera images were preferred, but the missing third dimension caused challenges for fast and precise grasping. Fusing RGB-D data from different perspectives could overcome this issue.

Lessons Learned and Conclusions

During the CENTAURO project and its evaluation, systematic integration has played a key role. Complex subsystems, such as the Centauro robot, the telepresence suit, and autonomous control functionalities, require extensive testing of isolated functionalities. We have learned not to underestimate the preparation effort a systematic evaluation requires. Misunderstandings between the operator and the referee occurred since details of individual task objectives were not communicated precisely.

In summary, we presented the final state of the CENTAURO system consisting of the flexible Centauro robot and a set of operator interfaces including a telepresence suit. We demonstrated the system capabilities for solving challenging, realistic, and wide-ranging disaster-response tasks. Our set of operator interfaces was even capable of coping with unexpected events, such as broken cameras or the temporal unavailability of individual interfaces. We further showed that neither pure teleoperation nor pure autonomy is desirable for controlling a complex robot for such challenging tasks. Instead, a set of control interfaces that address different tasks and support the operation on different levels of autonomy provides intuitive control and flexibility.

Nevertheless, the system speed to solve tasks is still significantly slower compared to the speed a human would require. We observed that control interfaces with a high degree of autonomy are in general faster in relation to the task complexity. Hence, extending the capabilities of autonomous functionalities is promising but requires considerable development effort. Regarding the speed of direct teleoperation through the telepresence suit, we discovered that force feedback is valuable as are other feedback modalities, such as vision. We identified a lack of system understanding as the main reason for slower operation speed.

In general, force feedback was found to be valuable to the operator and essential in certain tasks. However, it required noticeable hardware complexity of the telepresence suit. In future system development, a better tradeoff between the performance of the haptic feedback and the complexity of the hardware should be found by exploring different approaches.

Finally, given KHG's extensive experience in challenges of disaster-response systems, CENTAURO has overcome several of the described issues in demonstrating an effective locomotion approach, bimanual manipulation, and operator interfaces that reduce the cognitive load. The

CENTAURO system constitutes a significant step in enhancing safety for rescue workers in real-world disaster-response tasks.

Acknowledgment

This work was supported by the European Union's Horizon 2020 Programme under grant agreement 644839 (CENTAURO).

References

[1] K. Hiraoka, S. Tateishi, and K. Mori, "Review of health issues of workers engaged in operations related to the accident at the Fukushima Daiichi nuclear power plant," *J. Occupat. Health*, vol. 57, no. 6, pp. 497–512, 2015.

[2] M. Zucker et al. "A general-purpose system for teleoperation of the DRC-HUBO humanoid robot," *J. Field Robotics (JFR)*, vol. 32, no. 3, pp. 336–351, 2015.

[3] M. Schwarz et al. "NimbRo Rescue: Solving disaster-response tasks with the mobile manipulation robot Momaro," *J. Field Robotics (JFR)*, vol. 34, no. 2, pp. 400–425, 2017.

[4] A. Stentz et al. "CHIMP, the CMU highly intelligent mobile platform," *J. Field Robotics (JFR)*, vol. 32, no. 2, pp. 209–228, 2015.

[5] P. Hebert et al. "Mobile manipulation and mobility as manipulation—Design and algorithms of RoboSimian," *J. Field Robotics (JFR)*, vol. 32, no. 2, pp. 255–274, 2015.

[6] MotivRobotics. Accessed on: 2019. [Online]. Available: http://www.motivrobotics.com

[7] BostonDynamics, "Boston Dynamics: Changing your idea of what robots can do." Accessed on: 2019. [Online]. Available: https://www.bostondynamics.com

[8] Toyota, "Toyota unveils third generation humanoid robot T-HR3." Nov. 21, 2017. [Online]. Available: https://global.toyota/en/detail/19666346

[9] F. Negrello et al. "Humanoids at work: The WALK-MAN robot in a postearthquake scenario," *Robotics Automation Mag. (RAM)*, vol. 25, no. 3, pp. 8–22, 2018.

[10] T. Klamt et al. "Remote mobile manipulation through a centaur-like robot: Full-body telepresence and autonomous operator assistance," *J. Field Robotics (JFR)*, vol. 36, pp. 1–31, July 2019.

[11] Z. Ren, N. Kashiri, C. Zhou, and N. Tsagarakis, "HERI II: A robust and flexible robotic hand based on modular finger design and under actuation principles," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2018, pp. 1449–1455.

[12] N. Kashiri et al. "CENTAURO: A hybrid locomotion and high power resilient manipulation platform," *IEEE Robotics and Automation Letters (RA-L)*, vol. 4, no. 2, pp. 1595–1602, 2019.

[13] L. Muratore, A. Laurenzi, E. Hoffman, A. Rocchi, D. G. Caldwell, and N. G. Tsagarakis, "On the design and evaluation of XBotCore, a cross-robot real-time software framework," *J. Software Eng. Robotics (JOSER)*, vol. 8, pp. 164–170, 2017.

[14] A. Laurenzi, E. M. Hoffman, L. Muratore, and N. G. Tsagarakis, "CartesI/O: A ROS based real-time capable Cartesian control framework," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2019. doi: 10.1109/ICRA.2019.8794464.

[15] M. Mallwitz, N. Will, J. Teiwes, and E. A. Kirchner, "The Capio active upper body exoskeleton and its application for teleoperation," in *Proc. 13th ESA/Estec Symp. Advanced Space Technologies in Robotics and Automation (ASTRA)*, 2015. [Online]. Available: http://robotics.estec.esa.int/ASTRA/Astra2015/Papers/Session%204A/95936_Mallwitz.pdf

[16] D. Buongiorno, E. Sotgiu, D. Leonardis, S. Marcheschi, M. Solazzi, and A. Frisoli, "WRES: A novel 3DoF wrist exoskeleton with tendon-driven differential transmission for neuro-rehabilitation and teleoperation," *IEEE Robotics and Automation Letters (RA-L)*, vol. 3, no. 3, pp. 2152–2159, 2018.

[17] M. Gabardi, M. Solazzi, D. Leonardis, and A. Frisoli, "Design and evaluation of a novel 5 DoF underactuated thumb-exoskeleton," *IEEE Robotics and Automation Letters (RA-L)*, vol. 3, no. 3, pp. 2322–2329, 2018.

[18] D. Buongiorno, D. Chiaradia, S. Marcheschi, M. Solazzi, and A. Frisoli, "Multi-DoFs exoskeleton-based bilateral teleoperation with the time domain passivity approach," *Robotica*, pp. 1–22, 2019.

[19] T. Cichon, M. Priggemeyer, and J. Rossmann, "Simulation-based control and simulation-based support in erobotics applications," *Appl. Mechanics Materials*, vol. 840, pp. 74–81, 2016.

[20] T. Cichon, C. Schlette, and J. Roßmann, "Towards a 3D simulation-based operator interface for teleoperated robots in disaster scenarios," in *Proc. IEEE Int. Symp. Safety, Security, and Rescue Robotics (SSRR)*, 2016, pp. 264–269.

[21] T. Cichon and J. Roßmann, "Digital twins: Assist and support cooperation in human–robot teams," in *Proc. Int. Conf. Control, Automation, Robotics and Vision (ICARCV)*, 2018, pp. 486–491.

[22] T. Klamt and S. Behnke, "Anytime hybrid driving-stepping locomotion planning," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2017, pp. 4444–4451.

[23] M. Cordts et al., "The cityscapes dataset for semantic urban scene understanding," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 3213–3223.

[24] T. Klamt and S. Behnke, "Towards learning abstract representations for locomotion planning in high-dimensional state spaces," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2019. doi: 10.1109/ICRA.2019.8794144.

[25] D. Rodriguez, C. Cogswell, S. Koo, and S. Behnke, "Transferring grasping skills to novel instances by latent space non-rigid registration," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2018, pp. 1–8.

[26] D. Rodriguez and S. Behnke, "Transferring category-based functional grasping skills by latent space non-rigid registration," *IEEE Robotics and Automation Letters (RA-L)*, vol. 3, no. 3, pp. 2662–2669, 2018.

[27] D. Pavlichenko, D. Rodriguez, M. Schwarz, C. Lenz, A. S. Periyasamy, and S. Behnke, "Autonomous dual-arm manipulation of familiar objects," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots (Humanoids)*, 2018, pp. 1–9.

Tobias Klamt, Autonomous Intelligent Systems, University of Bonn, Germany. Email: klamt@ais.uni-bonn.de.

Diego Rodriguez, Autonomous Intelligent Systems, University of Bonn, Germany. Email: rodriguez@ais.uni-bonn.de.

Lorenzo Baccelliere, Humanoids and Human-Centred Mechatronics, Istituto Italiano di Tecnologia, Genoa, Italy. Email: Lorenzo.baccelliere@iit.it.

Xi Chen, Department of Robotics, Perception, and Learning, KTH Royal Institute of Technology, Stockholm, Sweden. Email: xi8@kth.se.

Domenico Chiaradia, Laboratory of Perceptual Robotics, Institute of Communication, Information, and Perceptual Technologies, Sant'Anna School of Advanced Studies, Pisa, Italy. Email: domenico.chiaradia@santannapisa.it.

Torben Cichon, Institute for Man–Machine Interaction, Rheinisch–Westfälische Technische Hochschule Aachen, Germany. Email: cichon@mmi.rwth-aachen.de.

Massimiliano Gabardi, Laboratory of Perceptual Robotics, Institute of Communication, Information, and Perceptual Technologies, Sant'Anna School of Advanced Studies, Pisa, Italy. Email: massimiliano.gabardi@santannapisa.it.

Paolo Guria, Humanoids and Human-Centred Mechatronics, Istituto Italiano di Tecnologia, Genoa, Italy. Email: paolo.guria@iit.it.

Karl Holmquist, Computer Vision Laboratory, Linköping University, Sweden. Email: karl.holmquist@liu.se.

Malgorzata Kamedula, Humanoids and Human-Centred Mechatronics, Istituto Italiano di Tecnologia, Genoa, Italy. Email: malgorzata.kamedula@gmail.com.

Hakan Karaoguz, Department of Robotics, Perception, and Learning, KTH Royal Institute of Technology, Stockholm, Sweden. Email: hkarao@kth.se.

Navvab Kashiri, Humanoids and Human-Centred Mechatronics, Istituto Italiano di Tecnologia, Genoa, Italy. Email: navvab.kashiri@iit.it.

Arturo Laurenzi, Humanoids and Human-Centred Mechatronics, Istituto Italiano di Tecnologia, Genoa, Italy. Email: arturo.laurenzi@iit.it.

Christian Lenz, Autonomous Intelligent Systems, University of Bonn, Germany. Email: lenz@ais.uni-bonn.de.

Daniele Leonardis, Laboratory of Perceptual Robotics, Institute of Communication, Information and Perceptual Technologies, Sant'Anna School of Advanced Studies, Pisa, Italy. Email: daniele.leonardis@santannapisa.it.

Enrico Mingo Hoffman, Humanoids and Human-Centred Mechatronics, Istituto Italiano di Tecnologia, Genoa, Italy. Email: enrico.mingo@iit.it.

Luca Muratore, Humanoids and Human-Centred Mechatronics, Istituto Italiano di Tecnologia, Genoa, Italy, and School of Electrical and Electronic Engineering, University of Manchester, United Kingdom. Email: luca.muratore@iit.it.

Dmytro Pavlichenko, Autonomous Intelligent Systems, University of Bonn, Germany. Email: pavlichenko@ais.uni-bonn.de.

Francesco Porcini, Laboratory of Perceptual Robotics, Institute of Communication, Information and Perceptual Technologies, Sant'Anna School of Advanced Studies, Pisa, Italy. Email: francesco.porcinia@santannapisa.it.

Zeyu Ren, Humanoids and Human-Centred Mechatronics, Istituto Italiano di Tecnologia, Genoa, Italy. Email: zeyu.ren@iit.it.

Fabian Schilling, Laboratory of Intelligent Systems, Swiss Federal Institute of Technology, Lausanne, Switzerland. Email: fabian.schilling@epfl.ch.

Max Schwarz, Autonomous Intelligent Systems, University of Bonn, Germany. Email: schwarz@ais.uni-bonn.de.

Massimiliano Solazzi, Laboratory of Perceptual Robotics, Institute of Communication, Information and Perceptual Technologies, Sant'Anna School of Advanced Studies, Pisa, Italy. Email: massimiliano.solazzi@santannapisa.it.

Michael Felsberg, Computer Vision Laboratory, Linköping University, Sweden. Email: michael.felsberg@liu.se.

Antonio Frisoli, Laboratory of Perceptual Robotics, Institute of Communication, Information and Perceptual Technologies, Sant'Anna School of Advanced Studies, Pisa, Italy. Email: antonio.frisoli@@santannapisa.it.

Michael Gustmann, Kerntechnische Hilfsdienst, Karlsruhe, Germany. Email: m.gustmann@khgmbh.de.

Patric Jensfelt, Department of Robotics, Perception and Learning, KTH Royal Institute of Technology, Stockholm, Sweden. Email: patric@kth.se.

Klas Nordberg, Computer Vision Laboratory, Linköping University, Sweden. Email: klas.nordberg@liu.se.

Jürgen Roßmann, Institute for Man-Machine Interaction, Rheinisch–Westfälische Technische Hochschule Aachen, Germany. Email: rossmann@mmi.rwth-aachen.de.

Uwe Süss, Kerntechnische Hilfsdienst, Karlsruhe, Germany. Email: u.suess@khgmbh.de.

Nikos G. Tsagarakis, Humanoids and Human-Centred Mechatronics, Istituto Italiano di Tecnologia, Genoa, Italy. Email: nikos.tsagarakis@iit.it.

Sven Behnke, Autonomous Intelligent Systems, University of Bonn, Germany. Email: behnke@cs.uni-bonn.de.