Interaction-Safe Humanoid Robot for Augmented Reality-Assisted Telemanipulation

Shuangyue Yu[1], Ran Duan[1], Haotian Cui[1], Tzu-Hao Huang[1], Chen Feng[2], Yingli Tian[3], Ziming Zhu[1], Hao Su[1,*]

[1] Biomechatronics and Intelligent Robotics Lab, City University of New York, City College, NY, United States
[2] Mitsubishi Electric Research Laboratories (MERL), MA, United States
[3] Department of Electrical Engineering, City University of New York, City College, NY, United States
* Corresponding author: hao.su@ccny.cuny.edu

ABSTRACT

Master-slave telemanipulators are a critical instrument for the operations of the hot cell and the operators manually control the master robot in front of the hot cell with potential safety hazards. In the glovebox operation phase, the operators manually handle hazard material with similar potential safety concerns. This paper proposes a holistic solution with hardware and software safety intelligence allows teleoperation from a safe control room to remotely control the operations in both hot cell and glovebox environments.

In terms of the hardware safety intelligence, the operators who wear the master whole-body control suit can safely work in a control room by using the body gesture to intuitively teleoperate the wheeled humanoid to manipulate the existing slave manipulator. In terms of software safety intelligence, an augmented reality (AR) system that is based on the fastest plane extraction algorithm we recently proposed is displayed to the operator with the potential to increase the task execution speed and reduce the task complexity for the operators. This AR system provides cognitive augmentation to achieve intuitive and hazard-free operations to enhance safety and efficiency. The proposed system is compatible with hot cells with different sizes and dimensions and can be implemented without any renovation of the existing infrastructure.

INTRODUCTION

The operations of the hot cell are a vital procedure in the nuclear industry. Nowadays, the operations of the hot cells mainly rely on the master-slave robot with a screen, where the operators stand in front of the hot cell to control the master robot manually. This kind of system has safety concerns that the potential radiation or contamination leaks threaten the safety of the first-line workers and the operator's safety has always been paramount in hot cell manipulations. For the protection of the operators, sophisticated structures of the hot cell are designed and heavy concrete walls are built. Safety concerns also force the company to put a lot of efforts on shielding and limiting the operation time for the crew. As technology progresses, the remote operation system of the hot cell is highly suggested in the nuclear industry. Besides, the master robot in the current operating system of the hot cell is controlled by the visual feedback of human. The operators still have to see the status of the slave robot inside of the hot cell and the lead glass is most commonly used to view the hot cell which is restricted by the limited field of view and imprecise distance estimation by human eye, therefore, it brings a bunch of problems in manipulating the slave robot. More importantly, the operators may still take the risk of exposing to contamination or radiation as they are working in front of the lead glass due to possible happened misoperation and radiation leakage. The camera system in hot cell empowers the remote control of hot cell operations. However, it still suffers from the narrow field of view and distance ambiguity due to a single camera. Putting multiple cameras inside the hot cell is not a realistic solution due to the limited space of the hot cell. And it is also a challenging problem to synthesize the information of one camera view with the other.

With the aforementioned problems in mind, a solution using a master-slave teleoperation system to substitute the first-line workers by the robot can perfect resolve the problem, as shown in Fig. 1. There are

some of the teleoperation robotic systems, RoboSally and Johnny-T, developed in the Applied Physics Laboratory of Johns Hopkins University [1,2]. Padir et al. also successfully handled high-consequence materials in gloveboxes with a ready-made full-body humanoid robot system (Valkyrie, NASA Johnson Space Center) [3]. But their applications are mainly focused on disaster response, safety inspection, and space tasks and their physical size are not fit for hot cell control.

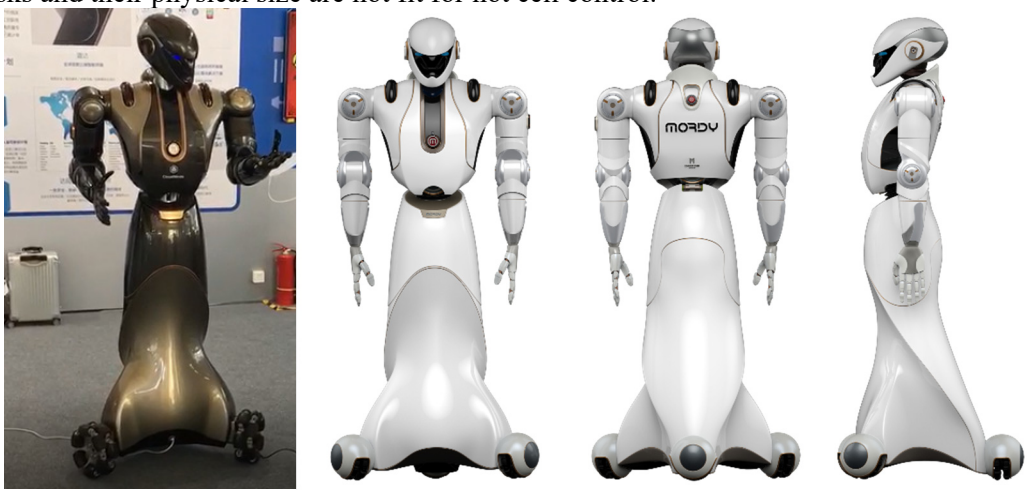

Fig. 1. Tele-manipulation humanoid robot system allows the operator to work remotely in a safe and hazard-free environment.

In this paper, a master-slave teleoperation system has been designed to manipulate the hot cells, which can greatly help to provide the first-line workers with maximum safety during hot cell manipulation and also increase the operational reliability of the master-slave robot systems. As using this system, the structure of hot cells does not need to change and it also has a good compatibility with the size of different hot cells. There are three major contributions in this work:

- The teleoperation humanoid system for hot cell manipulation offers inherently safe work condition for the hot cell operators. Our system consists of an environment-safe slave robot platform and a whole-body master control system. The environment-safe slave robot platform integrates human upper-body manipulation module, mobility module, and vision module and the whole-body master control system integrates wearable human motion capture module and a visual display module. The operator can manipulate the slave robot platform working in front of hot cells through the wearable module and real-time visual display information. In addition, the hot cell inner view will be acquired by an RGBD camera mounted on the slave robot arm inside the hot cell. Both the remote-control robot platform and camera system inside the hot cell can be disposed of as low-level radioactive waste.
- The augmented reality assisted vision system that has the potential for the task execution improvement and task error reduction. AR technology has revolutionized the way of many industrial and medical workflows [4, 5], where the sophisticated and precise manual operations are required. AR system could visualize a lot of useful information by blending digital graph and text to real-world view. This unlocks new possibilities for the way that the operator can be assisted or guided accurately by vision algorithm throughout the procedure [6]. It could also be used for low-cost training purpose, which will dramatically improve crews' performance [7].
- The real-time responded system was proposed by a fast and accurate plane extraction algorithm [8] that allows our system to estimate the robot pose and perform a precise control with a minor delay in visual feedback. Besides, the whole system also has a potential to manipulate other devices, such as gloveboxes, and can also realize more comprehensive application in the nuclear waste

management process.

ARCHITECTURE OF THE MASTER-SLAVE TELEOPERATION HUMANOID SYSTEM

As shown in Fig. 2, our overarching goal of this project is to build a custom-designed master-slave teleoperation humanoid system, which can substitute the first-line worker to manipulate hot cells. The system consists of two parts: environment-safe slave robot platform and whole-body master control system.

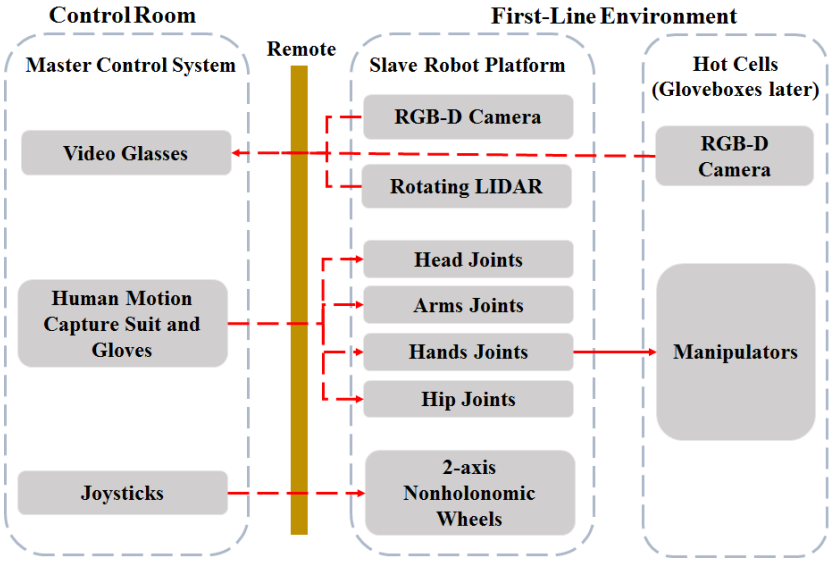

Fig. 2. Overview of the master-slave teleoperation system consisting of a portable motion capture suit to control the wheeled humanoid robot.

The environment-safe slave robot platform Mordy mainly consists of a bimanual industrial humanoid manipulator with multi-degrees of freedom (DOF), a three-axis movable wheeled chassis and vision acquisition system; The whole-body master control system mainly consists of a human motion capture module (suit and gloves) and video glasses module. By wearing the whole-body master control system, the operator can teleoperate the environment-safe slave robot platform to move freely in the working environment, access target, and control the bimanual homogeneous manipulator based on his upper-limbs and hands movement. In the meanwhile, the system uses an augmentation reality (AR) process method to process the original vision information which is acquired by RGBD cameras set on both robot platform and working environment and high accuracy LIDAR. As a control feedback, the processed image data will be displayed on the wearable video glasses in the real time. It can help the operator real-time adjust control strategy and movement of the slave robot platform.

Using the designed system, the operator can work safely and teleoperate the slave robot platform in a control room which located sufficiently far away from the hazardous work environment. The slave humanoid robot can substitute human being to work in nuclear facilities and manipulate hot-cells and gloveboxes.

For the realization of above-described system, the following design requirements have been carefully considered: 1) The slave robot platform needs to be able to work very flexible and dexterously, its bimanual manipulators need to have multiple DOF with high control accuracy, and the hands and fingers as the end effector need to be very dexterously and can perform various movements well. Besides, a safe-control strategy is necessary to avoid that the robot collides with the environment. 2) It is necessary for the operator to control the robot simply, directly and labor-saving. The continuous operation does not make the operator too tired, which will help to reduce the control error rate. 3) During teleoperation of the slave robot platform, the operator's vision needs to have a strong perception of the environment. As humans move in this

environment, our peripheral vision needs to provide a broader view than most of the cameras on the robot system. Therefore, through the image processing and augmented reality technology can help to meet this requirement.

Environment-Safe Slave Robot Platform

The environment-safe slave robot platform, Mordy, shown in Fig. 3. It is a 1.8 meters high upper-body humanoid robot with 34 DOF. It mainly consists of manipulation, mobile and vision module. The manipulation module consists of 3 major mechanical sub-assemblies: two arms and a torso. The arms consist of a total of 7 DOF arranged in a 3 DOF shoulder, 1 DOF elbow and 3 DOF wrist. The movement of 7 DOF redundant arms is similar to the arm movement of a human, which can provide flexibility in accomplishing tasks. Each arm is equipped with an 8 DOF cable-driven humanoid hand consisting of a 3 DOF thumb, 2 DOF index finger and three 1 DOF other fingers. The human thumb and index fingers should be dexterous enough to play a key role in the manipulation tasks, especially in substituting human finger to work in the gloveboxes. Thus, in order to ensure the flexibility of manipulator movement, the thumb can provide opposition, abduction-adduction, and flexion-extension movement, the index finger can provide abduction-adduction and flexion-extension movement, and other three fingers can provide flexion-extension movement. For a tactile sensing, an array of pressure sensors encased in a custom-made silica gel filler between finger skeleton. There is no direct force sensing available on the humanoid hand, so the control of fingers' actuator is mainly position based. The torso consists of a total of 4 DOF arranged in a 2 DOF neck and 2 DOF hip, which can help the head and upper-body easily turn around. The range of motion of each joint is shown in TABLE 1. Each DOF, even include fingers', is implemented by a series of custom-designed high torque density actuator, which can detect torque exerted by the external environment through current variation and realize flexible joint protection. More details of the actuators are presented in the following article.

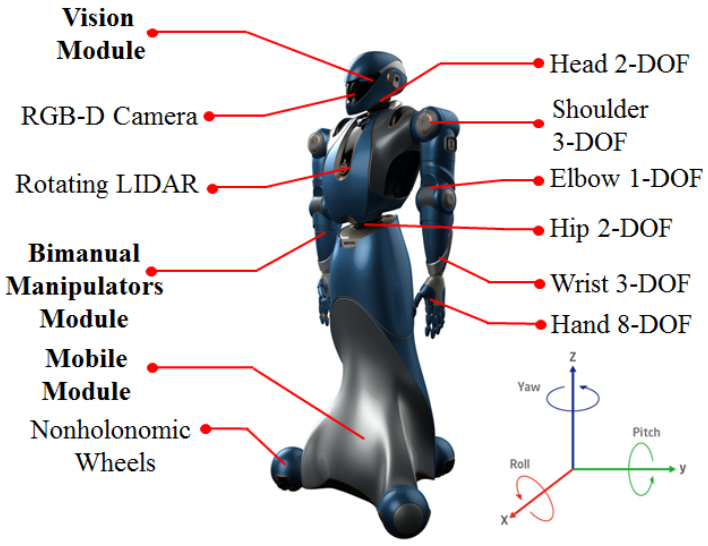

Fig. 3. The environment-safe slave robot platform Mordy consists of manipulation, mobility and vision module. Among them, the manipulation module totally has 34-DOF, includes two 7-DOF arms, two 8-DOF hands, 2-DOF head and 2-DOF hip joint; vision module includes an RGB-D camera and rotating LIDAR sensor; Mobile module uses three 2-axis nonholonomic wheels and can move freely in the indoor ground plane.

The mobility module mainly consists of a three-axis movable wheeled chassis. The angle between any two wheels is 120 degrees. The wheels are separately actuated by a Maxon DC motor. Through the coordination

movement of three wheels, the robot platform can move freely in the two-dimensional ground plane.
The vision module mainly consists of an RGB-D camera and high accuracy rotating LIDAR. An integrated FPGA in the system processes the RGB-D camera data and provides registered point clouds from the camera and LIDAR data. The integrated fully-dimmable LEDs can be used to illuminate the scene and work as an interface to reflect the running state of the robot. In addition, the LIDAR rotation speed can be varied, trading off denser point clouds for longer scan times. Besides, a wireless image transmission model is used for the acquisition of the hot cell inner camera system.

TABLE 1. Range of motion of joint

Joint Name	Motion	Range (degrees)
Head	Yaw	-120 ~ 120
	Pitch	-30 ~17
Arm-Shoulder	Pitch	-180 ~ 60
	Roll	0 ~ 150
	Yaw	-120 ~ 120
Arm-Elbow	Roll	-120 ~ 20
Arm-Wrist	Yaw	-120 ~ 120
	Pitch	-30 ~ 30
	Roll	-180 ~ 180
Leg-Hip	Yaw	-45 ~ 45
	Pitch	-29 ~ 17
Hand-Thumb Finger	Flexion/Extension, Abduction/Adduction & Rotation	
Hand-Index Finger	Flexion/Extension & Abduction/Adduction	
Hand-Others Finger	Flexion-Extension	

High Torque Density Actuator

As shown in Fig 4, the actuator is composed of the outer-rotor brushless motor, harmonic gearbox, magnetic absolute position encoder, and driver-control system. Different size actuator has different output peak torque, such as the smallest one is only about 10mm diameter with 0.5Nm output peak torque which can be used as finger actuator; the biggest one is about 80mm diameter with 180Nm output peak torque which can be used as a shoulder pitch joint actuator. Besides, the low-level control command is embedded in the driver-control system and can realize basic control methods, such as position and current control. High-level control devices can send a command to read and write the actuators' real-time information through a CAN bus communication protocol.

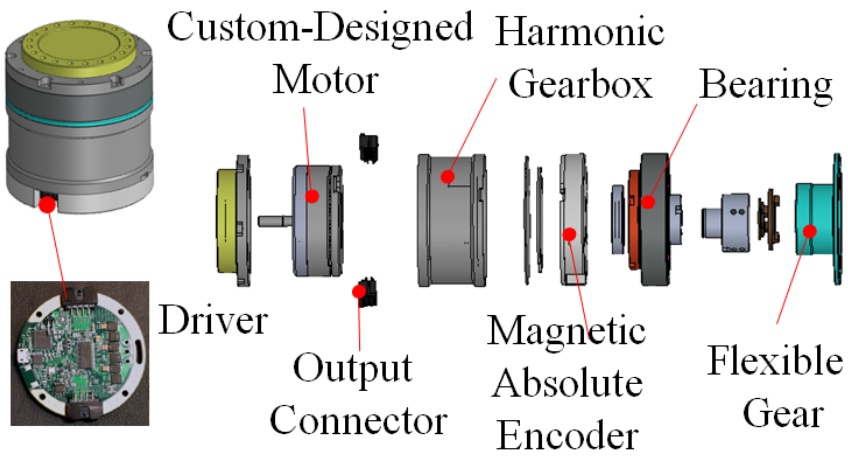

Fig. 4. A custom-designed high torque density actuator is composed of a custom-designed outer-rotor brushless motor, harmonic gearbox, magnetic absolute position encoder and driver-control system. There are different sizes and types of designed motors with different output torque. Besides, the low-level control command is embedded in the driver-control system and can realize basic control methods, such as position and current control. High-level control devices can send a command to read and write the real-time information actuator through CAN bus communication protocol by connecting the output connector.

TABLE 2 shows the performance comparison between our model (MSC-70-H50), HEBI-Robotics (X8-9, Biorobotics Lab, Carnegie Mellon University) and Anydrive (Robotic Systems Lab, ETH Zurich). Comparing with X8-9, the designed actuation system has better performance in output torque, respond speed, control accuracy, and even efficiency. Comparing with Anydrive, it has smaller size and the weight of our model is less than half weight of it.

TABLE 2. Our design electric actuator compared with HEBI X8-9 and ETH Anydrive

Property	Our Motor MSC-70-H50	HEBI Robotics X8-9	ETH Anydrive
Mass (g)	450	480	1000
Size (mm)	ϕ70 × 74 (L)	110 (L) × 73(W) × 45 (H)	Φ95 × 90 (L)
Rated Torque (NM)	14	8	15
Peak Torque (NM)	45	20	40
Rated Speed (RPM)	120	30	114
Accuracy (arcmin)	1	15	2
Power Efficiency (%)	65	35	60

Whole-Body Master Control System

The whole-body master control system mainly consists of a human motion capture module, video glasses module and joysticks module. Human motion capture module consists of a wearable suit and data gloves and can acquire human motion data in the real time by using Inertial Measurement Unit (IMU) nodes. Each IMU node has 3-axis accelerometer, 3-axis magnetometer, and 3-axis gyroscope and can detect real time attitude of user position. The video glasses module is for the visual display and we currently use a virtual reality headset, such as the Oculus Rift. Joysticks module is for teleoperate mobile module of the slave robot platform.

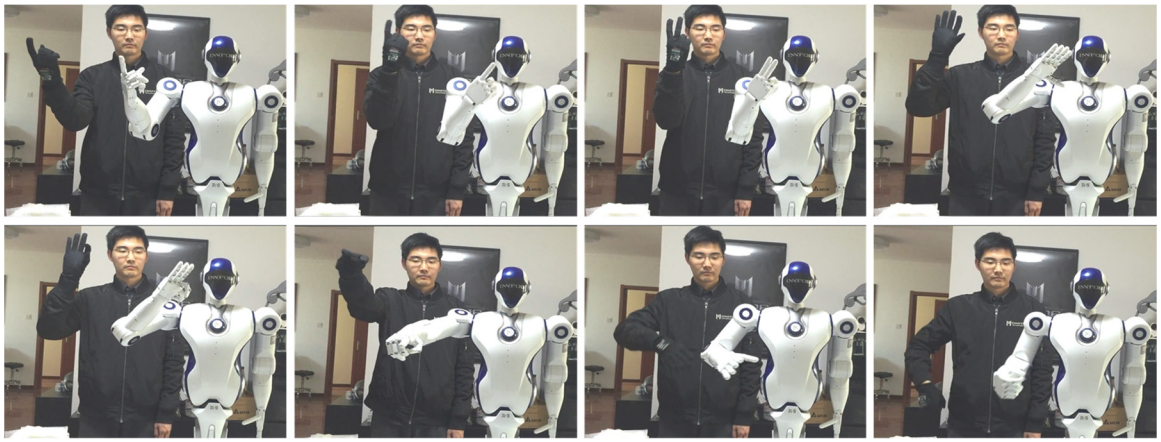

Fig. 5. By wearing the human motion capture module, the operator can control the slave robot platform through his/her body movement in the real time. The movement of the operator, even including finger movements, can be mirrored to the slave robot platform, and realize operator's direct control in a human-in-the-loop way.

As shown in Fig. 5, by wearing the whole-body master control system, the operator can control the slave robot platform through his/her body movement. The movement of the operator, even including finger movements, can be mirrored to the slave robot platform in the real time, and realize operator's direct control. In the meanwhile, the video image information is processed by AR technology, and the video glasses can receive and display the real time processed video stream to the operator. Based on that, the operator can adjust movement and complete the operation task continuously and smoothly.

METHODOLOGY OF AUGMENTED REALITY

We proposed an augmented reality (AR) system for hot cell manipulation and preliminary results on fast yet accurate and robust plane extraction for real-time camera pose estimation, scene reconstruction and object segmentation.

Augmented reality of hot cell manipulation

We present an RGBD camera-based AR system which can be mounted on the slave robot arm. Our AR system reduces the operation complexity of hot cell workflows in many aspects. For instance, as the operator is moving objects to their target positions, AR system will label the distance and pose of the object to assist grasping. Compare with the distance estimation from the human eye, the measurements of the depth camera and vision algorithm are accurate and reliable. The distances can be plotted as 3D vectors which are intuitively readable for the operator. An optimized trajectory will also be plotted for movement guidance. If the object is occluded, the AR will project the virtual graph to help the operator to see through the obstacles. In addition, the AR system facilitates the collaboration between operators and other people. The experts and supervisors can monitor the operation via the shared view and annotate into it to guide the

operators. Another critical workflow that our AR system will enable is training the operators through virtual objects and tasks before a costly operation start. For decades, companies are spending a lot of time and money on this, and an operation failure during the training could cause a devastating result. The cost, as well as risk, could be significantly reduced if the training operation is conducted by AR system.

Fast plane extraction

In order to accomplish the goals aforementioned, we present a preliminary work, i.e., a fast plane extraction approach [5], for fast and accurate camera pose estimation, which is the fundamental problem of AR. The overview of this algorithm is shown in Fig. 6. The process in each frame consists of three phases. In the first phase, an initial graph is constructed by dividing the point clouds into non-overlapping groups in the image space. Then the coarse planes are extracted in the second phase. The nodes that belong to the same plane are systematically merged by performing an agglomerative hierarchical clustering (AHC) on the initial graph. Two steps are working iteratively in AHC to achieve the proper planes fitting until the mean squared error (MSE) exceeds a threshold: (1) finding the minimum plane fitting MSE region and (2) merging it with one of its neighbors which results in the minimum plane fitting MSE of the merged region. While coarse segmentations are already used in many applications, a pixel-wise region growing approach is employed to refine artifacts in the coarse segmentation in the last phase in order to obtain better results. The presented algorithm can reliably detect all major planes in the scene at a frame rate of more than 35Hz for 640×480 point clouds using an ordinary laptop with Intel Core i7-2760QM CPU of 2.4GHz and RAM of 8GB. The accuracy evaluation results on two SegComp [9] datasets, i.e., ABW and PERCEPTRON, are shown in Fig. 7 and TABLE 3, where the image datasets are acquired using ABW structured light scanner and a Perceptron laser rangefinder. The ground truth of each database is given by an average number of connected instances.

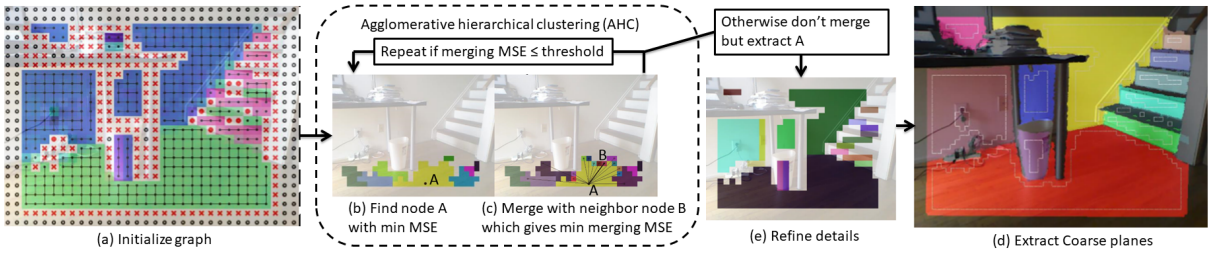

Fig. 6. Algorithm overview. (a) shows node (black dot) and edge (line) in the initialized graph, where red 'x', black 'o', and red dot showing node rejected by depth discontinuity, missing data, and too large plane fitting MSE, respectively. (b) and (c) show AHC process where graph nodes merged at least once. Black lines in (c) show all edges coming out from the node A. The node B that gives the minimum plane fitting MSE when merging the node A is marked by a thick line. (d) shows the extracted coarse planes using colored regions, which are finally refined in (e) if required by the application.

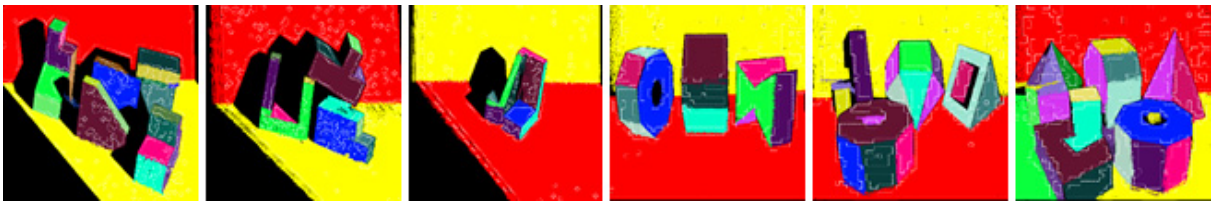

Fig. 7. 3D planes segmentation results on two SegComp dataset

TABLE 3. Accuracy evaluation on SegComp datasets

Property	ABW	PERCEPTRON
Regions in ground truth	15.2	14.6
Correctly detected	12.8 (84.2%)	8.9 (60.9)
Orientation deviation (°)	1.7	2.4
Under-segmented	0.0	0.2
Over-segmented	0.1	0.2
Noise (non-existent)	0.7	2.1
Missed (not detected)	2.4	5.1

Plane-based pose estimation

With the prior knowledge of the hot cell environment, i.e., the 3D model of the inner static scene of the hot cell, we can estimate the camera pose as well as objects' pose using plane-to-plane registration. The plane measurement of static model and camera current observation can be represented by plane parameters $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$, respectively. Let $\boldsymbol{\pi}_m = (\boldsymbol{n}_m^T, \boldsymbol{d}_m)^T$ and $\boldsymbol{\pi}'_m = (\boldsymbol{n}'^T_m, \boldsymbol{d}'_m)^T$ be the m-th corresponding pair of 3D planes, where the vector $\boldsymbol{n}$ and $\boldsymbol{d}$ denote the unit normal vector and the distance to the origin of the camera coordinate system, respectively. The rigid body transformation, i.e., rotation and translation $[\mathrm{R}, \mathbf{t}]$ can be computed by:

$$\underset{R}{\operatorname{arcmin}} \sum_i \|\boldsymbol{n}'_i - R\boldsymbol{n}_i\| \qquad \text{(Eq. 1)}$$

$$\boldsymbol{t} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b} \qquad \text{(Eq. 2)}$$

where matrix $A = (\boldsymbol{n}_1^T, \boldsymbol{n}_2^T, \dots)^T$ and $b = \big((\boldsymbol{d}_1 - \boldsymbol{d}'_1), (\boldsymbol{d}_2 - \boldsymbol{d}'_2), \dots\big)^T$. Thus, the absolute pose of camera is determined in the pre-defined coordinate system of the 3D model of the hot cell. This bridges the digital world to the real world and enables the AR system to project any virtual object to the view.

CONCLUSIONS

The proposed master-slave teleoperated humanoid robot system can be disposed of as low-level radioactive waste for remote hot cell manipulation. This system creates a risk-free environment for the first-line worker and assists the operator by using an AR system. We also demonstrated that the fast plane extraction algorithm and AR system is capable of real-time camera pose estimation and object segmentation. The minor delay of the visual feedback allows our system to perform the real-time control. The proposed system can be implemented for any hot cell equipment by simply mounting an RGBD camera on the slave robot inside the hot cell. It significantly increases the safety and control performance in the hot cell manipulation scenario. In the future, the whole system also has the potential to manipulate other devices, such as gloveboxes and can also realize more comprehensive application in the treatment of nuclear waste

management.

REFERENCES

1. E. Tunstel, K. Wolfe, M. Kutzer, M. Johannes, C. Brown, K. Katyal, M. Para, and M. Zeher, "Recent Enhancements to Mobile Bimanual Robotic Teleoperation with Insight Toward Improving Operator Control, " *Johns Hopkins APL Technical Digest*, vol. 32, no. 3, pp. 584-594 (2013).
2. K. Katyal, C. Brown, S. Hechtman, M. Para, T. McGee, K. Wolfe, R. Murphy, M. Kutzer, E. Tunstel, M. McLoughlin, and M. Johannes, "Approaches to Robotic Teleoperation in a Disaster Scenario: Form Supervised Autonomy to Direct Control, " *IEEE International Conference on Intelligent Robots and Systems*, pp. 1874-1881, (2014).
3. T. Padir, H. Yanco, and R. Platt, "Towards Cooperative Control of Humanoid Robots for Handling High Consequence Materials in Gloveboxes, " *Waste Management Conference*, March (2017).
4. S. Dong, A. H. Behzadan, F. Chen, and V. R. Kamat, "Collaborative Visualization of Engineering Processes Using Tabletop Augmented Reality, " *Advances in Engineering Software*, vol. 55, pp. 45-55 (2013).
5. L. Soler, S. Nicolau, J. Schmid, C. Koehl, J. Marescaux, X. Pennec, and N. Ayache, " Virtual Reality and Augmented Reality in Digestive Surgery, " *Proceeding of The 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, Arlington, VA, USA, Nov. 2 - 5, pp. 278-279, Nov (2004).
6. D. Teber, S. Guven, T. Simpfendörfer, M. Baumhauer, E. O. Güven, F. Yencilek, A. S. Gzen, and J. Rassweiler, "Augmented Reality: A New Tool to Improve Surgical Accuracy During Laparoscopic Partial Nephrectomy? Preliminary in Vitro and in Vivo Results, " *European urology*, vol. 52, no. 2, pp. 332-338 (2009).
7. De Crescenzio F., Fantini M., Persiani F., Di Stefano L., Azzari P., and Salti S, "Augmented Reality for Aircraft Maintenance Training and Operations Support, " *IEEE Computer Graphics and Applications*, vol. 31, no. 1, pp. 96-101 (2011).
8. C. Feng, Y. Taguchi, and V. R. Kamat, "Fast Plane Extraction in Organized Point Clouds Using Agglomerative Hierarchical Clustering, " *Proceedings of the IEEE Internatioanl Conference on Robotics and Automation (ICRA)*, Hong Kong, May 31-Jun 7, pp. 6218-6225, (2014).
9. A. Hoover, G. Jean-Baptiste, X. Jiang, P. J. Flynn, H. Bunke, D. Goldgof, K. Bowyer, D. Eggert, A. Fitzgibbon, and R. Fisher, " An Experimental Comparison of Range Image Segmentation Algorithms, " *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 7, pp. 673-689 (1996).