Augmented Teleoperation for D&D

Young Soo PARK[1], Joohee KIM[2], and Byung Seon CHOI[3]

1. Nuclear Engineering Division, Argonne National Laboratory, Lemont, IL, USA (ypark@anl.gov)
2. Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, IL, USA (joohee@ece.iit.edu)
3. Korea Atomic Energy Research Institute, Daejeon, Republic of Korea (bschoi@kaeri.re.kr)

Abstract: This paper presents a development of a new teleoperation method to enhance the performance of telerobotic operations for nuclear facility decontamination and decommissioning (D&D). It focuses on enhancement of human-robot interface based on introduction of virtual reality (VR) and augmented reality (AR) technologies which can provide immersive artificial environment for the operator to interact with. In this regard, the key technology innovation is realized by 3D sensing and reconstruction and 'virtual fixtures'. The development will allow to use simple and robust robotic systems for complex and dexterous D&D task operations.

Keyword: Augmented reality, Teleoperation, D&D

1 Introduction

Recently significant new advances in virtual reality (VR) and augmented reality (AR) technology have emerged. These technologies bring the human-robot interface to a new level, and potentially provide tools for enhancing the performance of remote operation. VR can provide immersive artificial environment for the operator to interact with. As an advanced concept, AR adds to it the blending or artificial contents with real world - users are able to interact with virtual contents in the real world.

An effective teleoperator interface does not have to be complicated, but it has to be done well. A key to success in implementation of such innovation in teleoperation is accurate telepresence. It has been known that information presented inaccurately in VR and AR media may confuse the operator, and may cause damage to the equipment and environment. To this end, this paper presents a series of innovations which makes VR an AR applicable for contact manipulation applications. Such innovations include accurate 3D sensing, reconstruction and dynamic tracking [1], multiple sensor calibration and coordinate mapping, multi-modal perceptual feedback [2], as well as robot control system integration.

In this regard, an enhanced teleoperation, namely augmented teleoperation, system is implemented, which includes a two-arm robot manipulator system, virtual environment for enhanced telepresence, and sensor-based augmented reality capability. The software and hardware integration is performed under a Linux-based distributed open architecture, namely Robot Operating System (ROS). The enhanced teleoperation is expected to improve the precision and efficiency of in-situ remote operation. Test operation is performed for a manipulation task representative of D&D operation.

<u>*System Overview*</u>

To develop and demonstrate the utility of augmented-reality technologies in implementing enhancement of teleoperation performance, teleoperation testbed is composed. As illustrated in Fig. 1, the augmented teleoperation system incorporates the capabilities of real-time 3D reconstruction, augmented reality, and teleoperation testbed.

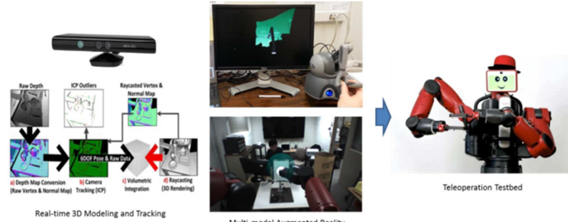


Fig.1. Augmented Teleoperation System Overview

2 Enhanced 3D Reconstruction

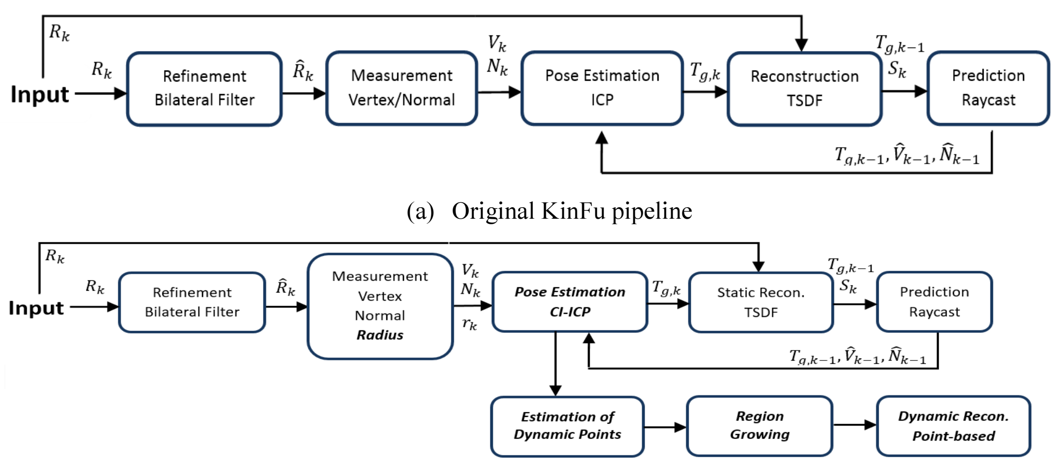
Nowadays, real-time 3D sensing and reconstruction technology has been established, which is capable of recreating a real scene within a virtual three-dimensional space. Recently, low-cost depth cameras such as the Microsoft Kinect [3] have been introduced, which are capable of generating depth maps at real-time rates. Several software options are available to communicate with the Kinect, for example Microsoft's KinectFusion SDK, or KinFu.

Our 3D reconstruction is based on KinFu API for 3D reconstruction. Fig. 2(a) illustrates the processing pipeline of the 3D reconstruction, which involves 1) depth map conversion, 2) camera tracking, 3) volumetric fusion, and 4) raycasting. Although the recent advances have made real-time 3D reconstruction possible, the state-of-the-art has some limitations that they are subject to pose estimation error, and are only capable of tracking relatively static objects. Thus modifications are made on the processing pipeline to overcome such limitations. Fig. 2(b) shows the enhanced processing pipeline with improved reconstruction accuracy and dynamic object tracking capabilities.

<u>Enhanced Static Reconstruction</u>

Pose estimation is at the heart of 3D reconstruction. The 3D sensors have been shown to have systematic errors in the depth measurement, which are well studied and modelled for Kinect sensors [4]. KinFu uses truncated signed distance function (TSDF) for 3D reconstruction, whereas the systematic error model is incorporated in the TSDF [5]. However, the systematic error models have never been incorporated into the pose estimation. Therefore, we have developed a method to incorporate the systematic error models into pose estimation process in the form of confidence indicator for each depth value.

The raw depth map from the depth sensor is bound to be noisy and usually contain holes. The multi-resolution anisotropic diffusion based depth refinement algorithm proposed in [6] uses both RGB and depth information to achieve hole-filling method in real time. This hole-filling approach was primarily used as depth map refinement stage for view synthesis. The depth value created for the holes have errors proportional to the nearest measured depth value. We designed a method to model this error as a confidence indicator (CI) for each depth value and use it in pose estimation process. Subsequently the CIs are used to weigh both in the pose estimation and reconstruction stages.

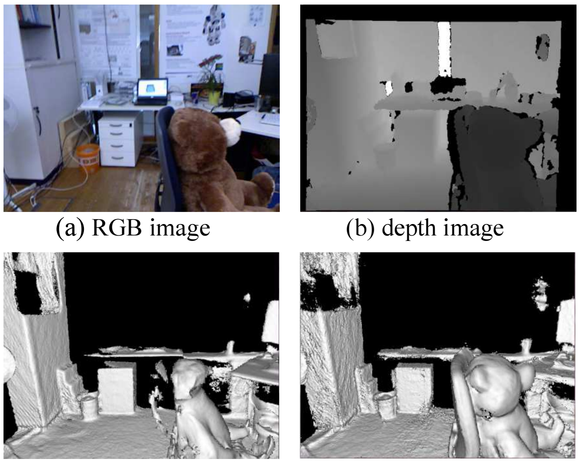


(a) Original KinFu pipeline

(b) Enhanced pipeline

Fig.2. Real-time 3D Reconstruction Pipeline

We have performed a series of test operations to present quantitative and subjective quality results demonstrating the advantages of using confidence indicator for each depth value from systematic errors and filtering errors in pose estimation and 3D reconstruction stage. Fig. 3 illustrates comparison of the 3D reconstruction results. The test results has shown that the confidence indicator based methods give 59.60% better pose estimate than baseline algorithm in quantitative measures and achieves a significant subjective quality improvement in 3D reconstruction.



(a) RGB image (b) depth image

(c) from original KinFu (d) using WICP with WTSDF

Fig.3. Comparison of 3D reconstruction of "Teddy" at the 400th frame

Dynamic Reconstruction

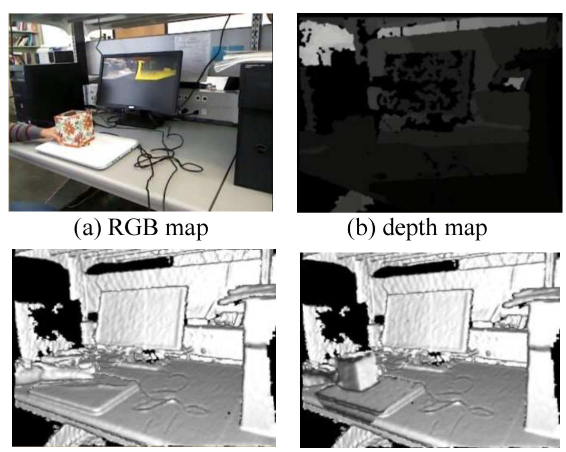
Another drawback of the KinFu pipeline is that if a relatively fast moving object is introduced in the scene, data association fails and there occur tracking errors in the camera pose estimation. To overcome this limitation, we have enhanced the 3D reconstruction for dynamic scene.

The enhanced method is based on the point-based method for dynamic tracking [7]. The approach discussed in [7] is very much similar to that of the conventional 3D reconstruction methods discussed earlier. 1) The data from the depth sensor is pre-processed; 2) the current 6 degree-of-freedom (6DoF) pose of sensor relative to the scene is estimated; and 3) the estimated pose is used to convert depth samples into a unified coordinate space and fuse them into an accumulated global model. The only difference is that this method uses point-based representation throughout the reconstruction process. The point-based fusion method works without the overhead of converting between representations. Along with that, it introduces the use of radius map [8] and confidence counter in the ICP algorithm to detect the dynamic candidates from the non-corresponding points. These dynamic candidates are then segmented using a hierarchical region growing method and used for scene reconstruction. It further uses the surface splatting method [9] for the final 3D reconstruction of the scene.

We have further improved the above point-based reconstruction method to enhance accuracy, reliability and speed. In this regard, the following enhancements are made:

- extended the radius map and confidence indicator to three dimensions,
- combined the point-based fusion and volumetric representation to maintain the level of the reconstruction quality, while reducing the processing time and memory usage,
- added the normal map as a similarity attribute to improve the region growing method,
- implemented it on the GPU to maintain its processing speed,
- the dynamic parts are added into the global model in real-time.

The new algorithm was tested on a reference video set in comparison with the baseline method. The new approach was three times faster, 52% better in absolute tracking accuracy, and resulted in better dynamic tracking as illustrated in Fig. 4.



(a) RGB map (b) depth map

(c) results of KinFu (d) proposed method

Fig.4. Comparison between the original KinFu and our proposed method for moving_box_1 scene

3 Augmented Teleoperation

This section presents implementation of an augmented teleoperation system, which is an enhanced teleoperation concept via use of virtual fixtures. The concept of virtual fixtures - an artificially generated geometric surface overlaid on human perception - was first introduced to enhance teleoperation performance [10, 11]. Such guidance is expected to reduce the operator's mental burden, facilitate precision motion, and improve stability. Also since it is based on local sensory feedback, it is applicable to operation of simple slave robots, i.e. those not requiring complex bilateral system. Fig. 5 conceptually illustrates the architecture of such an augmented reality system.

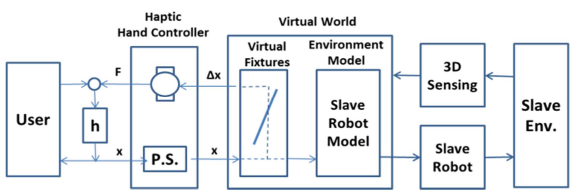


Fig.5. Augmented Teleoperation System Architecture

Teleoperation Testbed

The enhanced teleoperation method is implemented on the teleoperation test bed as illustrated in Fig. 6. In this system, a two-arm robot is used as the slave robot, which is remotely controlled by the human operator with a haptic device. A 3D sensor fixed on top of the robot is used to capture the 3D scene in front of the robot. The 3D geometry as well as the camera image of the environment is displayed on an immersive VR headset.

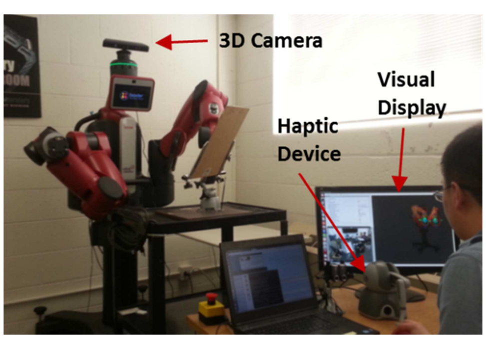


Fig.6. Augmented Teleoperation Testbed

A tele-operation software system is developed using the Robot Operating System (ROS) [12] and the RViz visualization environment [13] for operation of the whole system. ROS allows various utilities for robot control, and RViz allows visual display of sensor information and 3D task environment.

Implementation of Augmented Reality

Implementation of virtual fixtures requires accurate multi-modal augmentation in the teleoperator interface. It is accomplished by matching the visual and haptic rendering transformations as illustrated in Fig. 7.

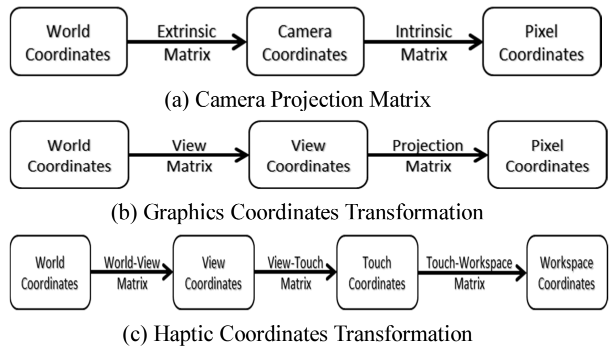


Fig.7. View Perspective Transformations for Visual-Haptic Rendering

Graphics Rendering

The graphics rendering module of the test bed software accomplishes the goals including (i) displaying the actual robot workspace, (ii) rendering a virtual cursor for the haptic stylus, reflecting its position and orientation, and (iii) rendering a set of artificial geometries as virtual fixtures to assist teleoperation. The software uses OpenGL and OpenCV for graphics rendering.

The live image captured by Baxter head camera shows the workspace in front of the robot. To achieve augmented reality, a virtual scene is established and displayed on top of the actual camera image. It should overlap with the actual scene with sufficient accuracy so that the Baxter arm, which is clearly in the actual scene, can interact with the virtual fixture, which is included in the virtual scene. OpenGL provides a series of matrix stacks to define the mapping from the virtual scene to the display interface. By comparing Fig. 7(a) with Fig. 7(b), it's clear that both the virtual and actual scenes will share pixel coordinates, and in turn be merged on display, if

- the world coordinates of the virtual scene are the same as that of the camera,
- view matrix equals to camera extrinsic matrix,
- projection matrix coincides with camera intrinsic matrix.

The view coordinates of the virtual scene and camera are matched according to the above conditions.

Haptic Rendering

In the teleoperation test bed, a haptic device is used by the operator to interact with the virtual scene. The haptics rendering module of the software is responsible of (i) map the haptic device to the actual scene, (ii) place the virtual fixture in the scene, and (iii) guide the movement of the robotic arm. The software uses OpenHaptics SDK for haptics rendering.

The haptic workspace is the physical space reachable by the haptic device. In the teleoperation test bed, the haptic workspace should be mapped to the intersection of the range of motion of the robot arm and the field of view of the head camera. In this way, the virtual fixture placed by the haptic device is always visible and reachable. To ensure the haptic feedback reflects the shape of the touchable objects, the mapping should be uniform, i.e. the scale factor should be the same for all dimensions.

OpenHaptics provides a series of matrix stacks to convert the coordinates of haptic workspace to world coordinates, as shown in Fig.7(c). World-View matrix defines the transformation to the camera coordinate frame. It is obtained from OpenGL. View-Touch matrix defines the rotation and translation of the haptic workspace relative to view coordinates independent of the workspace mapping. Touch-Workspace matrix defines the mapping of the workspace to view coordinates. The mapping contains a scale to map the workspace and a translation to orient the workspace to the target mapping in view coordinates.

Test Operation

Test teleoperation was performed where the task was to draw a circle following a circular pattern on a flat panel. Fig. 8 illustrates the overview of teleoperation process based on virtual fixtures.

- A live image captured by the Baxter head camera is displayed to the controller (Fig. 8(a)).
- A virtual fixture, a conical frustum, is introduced to the testbed (Fig. 8(b)).
- This virtual fixture is a haptic-enabled so that it is touchable by the haptic stylus. The haptic cursor stays on top of the surface when in contact with the virtual fixture (Fig. 8(c)).
- After the virtual fixture is placed, the haptic device starts guiding the movement of the Baxter arm (Fig. 8(d)). As the haptic cursor is constrained by the virtual surface, it provides guidance of the movement of robot arm.

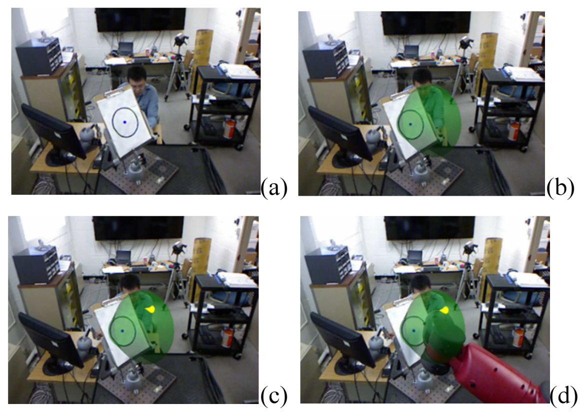
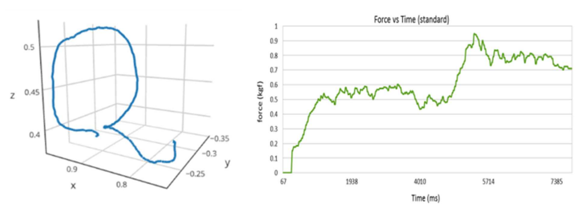


Fig.8. Guided Teleoperation by virtual fixture

During the test operation, the position and force at the robot hand was recorded, which can be obtained from the Baxter ROS package nodes. Fig. 9(a) shows the path of the robot hand, which approximately followed the circular pattern. It was particularly effective in maintaining the depth of the motion on the flat panel. Fig. 9(b) shows the magnitude of the aggregated contact force during the operation. It is noted that the contact force remained at relatively constant level during the kinesthetic interaction of following the circle on the panel.



(a) Positon (b) Force

Fig.9. Plot of the Position and Contact Force at the Robot Hand during the Test Operation

4 Conclusions

As an effort to fill the key technology gap required for practical remote system deployment for complex tasks in nuclear applications, a new concept of augmented remote operation method is introduced. This concept aims at augmenting virtual reality-based operator aid to simplify the remote operation task and improve task precision. The development has focus on enhancement of 3D sensing and reconstruction technology, and implementation of a telerobotic testbed based on an open source technology basis. The concept of virtual fixtures has been implemented in the ROS based test bed robot environment. The integrated technology will allow teleoperation of dexterous manipulation tasks using simple and rugged robot system suitable for D&D operations.

Acknowledgment

This work was supported by Korea Atomic Energy Research Institute as a part of a larger program by Ministry of Science, ICT, and Future Planning, South Korea.

References

[1] Y. S. Park, Z. Zhao, Z. Xia, J. Kim, “Accurate and Reliable 3D Sensing and Reconstruction for Teleoperation,” Proc. ANS Topical Meeting on D&D and Remote Systems, 18069, July 2016

[2] Y. S. Park, H. Kang, T. F. Ewing, et al, "Enhanced Teleoperation for D&D," Proc. IEEE ICRA (2004)

[3] J. Smisek, M. Jancosek, and T. Pajdla, “3D with Kinect”, in Proc. IEEE Intl. Conf. on Comp. Vision Workshops pages 1154-1160, Nov. 2011.

[4] A. Segal, D. Haehnel, and S. Thrun, “Generalized-ICP,” In Proceedings of Robotics: Science and Systems, Seattle, USA, 2009

[5] C. Kerl, J. Sturm, and D. Cremers, “Robust Odometry Estimation for RGB-D Cameras,” In Proc. IEEE Int’l Conf. Robotics and Automation (ICRA), May 2013

[6] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, “A benchmark for the evaluation of RGB-D SLAM systems,” in Proc. Int. Conf. Intell. Robot Syst., 2012, pp. 573–580.

[7] M. Keller, D. Lefloch, M. Lambers, S. Izadi, T. Weyrich and A. Kolb. “Real-time 3D Reconstruction in Dynamic Scenes using Point-based Fusion,” In 2013 International Conference on 3D Vision (3DV), pp.1-8, 2013.

[8] T. Weise, T. Wismer, B. Leibe and L. Van Gool. “In-hand scanning with online loop closure,” in Proc. IEEE International Conference Computer Vision Workshops, pp. 1630-1637, 2009.

[9] M. Zwicker, H. Pfister., J. V. Baar, and M. Gross. “Surface splatting,” in Procs. SIGGRAPH, pp. 371-378, 2001.

[10] L.B. Rosenberg, “Virtual fixtures: Perceptual tools for telerobotic manipulation,” Ph.D. Thesis, Dept. of Mechanical Engineering, Stanford University (1994)

[11] Y. S. Park, H. Kang, T. F. Ewing, et al, "Enhanced Teleoperation for D&D," Proc. IEEE ICRA (2004)

[12] http://wiki.ros.org

[13] http://wiki.ros.org/rviz