Proceedings of the European Nuclear Young Generation Forum
ENYGF'21
September 27-30, 2021, Tarragona, Spain

The Importance of Synthetic Viewing for Teleoperation Tasks in Hazardous Environments

Pacheco-Gutierrez, Salvador[1*], Caliskanelli, Ipek[1] and Skilton, Rob[1]

[1] UK Atomic Energy Authority – Remote Applications in Challenging Environments (UKAEA-RACE), United Kingdom

*Corresponding author: *salvador.pacheco-gutierrez@ukaea.uk*

I. INTRODUCTION

Nuclear power, conceived during the second world war as a mass destruction technology, has now become the primer technology aiming at low-carbon emission source of energy. Furthermore, the joining of the once divided West and East has led to a worldwide research ground that involves not only physics or chemistry but many other engineering fields, such as robotics, programming, and more recently artificial intelligence (AI) to make fusion energy a reliable source.

Recent advances in robotics and communications have allowed performing risky tasks such as handling, maintenance and decommissioning using teleoperation, applying unilateral or bilateral control. Unilateral control refers to a situation where one side follows the position or forces of the other. Typically, this would be a remote manipulator following a human-interface device (HID)[1]. Bilateral control is a scheme where both, HID and remote manipulator, follow each other in a closed loop. Moreover, teleoperation helps saving time, effort and most importantly lives.

With recent developments in graphics processing, together with highly accurate 3D sensing devices capable of providing high-quality point cloud[2] data at high frame rates, the creation of live virtual environments for aiding remote handling, maintenance and inspection is now a possibility at a cheaper cost. However, in a nuclear teleoperation scenario, the real-time transmission of large high-resolution point cloud data still represents a challenge due to bandwidth limitations combined with far geographical locations. To counteract this challenge, the information gathered can be processed using computer vision and deep learning algorithms to compress, segment, filter and effectively transmit only important information. Moreover, current development of 5G technologies appears to be promising in solving latency related problems [1].

This manuscript is structured as follows: Section II addresses the challenges in teleoperation with regards to its viewing systems. Section III discusses what virtual reality and synthetic viewing are, and how these technologies can be used to improve the user experience of remote operators in the nuclear industry. Lastly, in Section IV we discuss current work conducted in the UKAEA to address the challenges of integrating real-time synthetic viewing into teleoperated systems.

II. VISION SYSTEMS FOR ROBOTIC TELEOPERATION IN NUCLEAR

A fundamental aspect for successful teleoperation tasks is the knowledge of the inner state of the remote system and its surroundings in real-time; this is achieved using interoceptive and proprioceptive sensors, respectively. For the former, joint positions and velocities can be obtained through optical encoders attached to each joint. The latter lies in gathering real-time information using video and/or 3D cameras as well as any other external sensing device. As can be inferred, the accurate and fast transmission of this information is crucial for an adequate functioning of any teleoperation system. With focus on vision, we briefly describe a few of the systems utilised in nuclear applications, their importance, and current challenges associated with those systems.

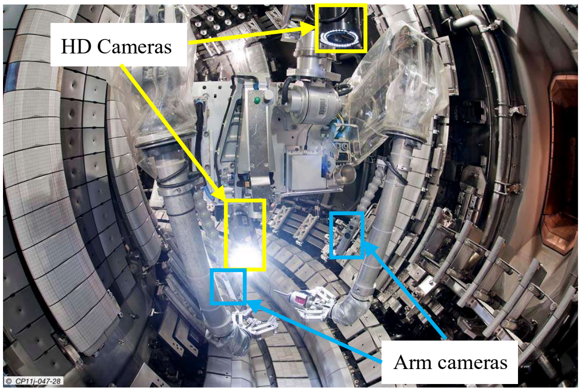


Figure 1. Remote Handling MASCOT Manipulator inside the JET Vacuum Vessel (May 2011).

In the case of fusion reactors, metallic tiles containing lithium are used, together with cooling systems and complex

[1] Due to the political incorrectness of the master-slave terminology, in this paper we refer human interface device (HID) to that is known as master, and the remote manipulator to what is refereed as the slave.

[2] A 3D point cloud is a set of points in $\mathbb{R}^3$ usually consisting of multiple attributes such as colour, texture, normal, etc.

magnetic coils, to contain the plasma within the vacuum vessel [2]. This results in radiation damage to the tiles that require adequate inspection and regular replacement (Figure 1). The maintenance activities are extremely dangerous for human beings due to the high radiation not only inside the vessel but on its near surroundings.

The robotic device shown in Figure 1 called MASCOT (Italian: MAnipulatore Servo COntrollato Transistorizzato) is a 1960's built teleoperated dual arm robotic manipulator with 7 degrees of freedom currently used for maintenance tasks inside the JET vessel using a bilateral control, more information about the remote handling system of JET can be found in [3]. The vision system of MASCOT consists of two high definition cameras mounted above and behind the operative area. Two extra cameras are mounted on each one of the arms assisting the operator with close-up views. This vision setup provides remote operators with live image/video feed of MASCOT's surroundings (Figure 2).

However, regardless of the quality of the cameras and the flexibility on its positioning, this solution fails to provide the operator with a sense of depth and volume (3D) in real-time. This is particularly important when performing grasping, bolting and other related tasks, and can drastically affect operators training time. A partial solution to this consists on integrating a virtual reality scenario together with the vision system (see Section III).

Figure 2. JET Remote Handling Control Room and MASCOT vision system.

Another application is the decommissioning of gloveboxes containing radioactive material. Although, usually they represent a low risk for operators, legacy or damaged gloveboxes can represent a high risk and a key challenge in reduction of hazardous waste. The need for a solution to this is such that recently Sellafield LTD granted a multimillionaire contract to a UK-based firm to tackle this problem [4]. Moreover, the National Nuclear Laboratory (NNL) recently released a UKRI funded challenge for the decommissioning of alpha-contaminated gloveboxes [5]. Such task consisted in utilising novel robotic-aided cutting technologies.

Depending on the levels of radiation present within the gloveboxes, some of these are simply stored in confined locations, thus creating more nuclear waste. As the need for decommissioning appears obvious, a reliable and robust vision system providing the operators with an adequate insight of the contents of the gloveboxes is required. Furthermore, due to the limited space and confined nature of the gloveboxes, the usage of multiple standard cameras, as in the case of JET, is not a feasible solution. Under this circumstance, the amount of visual information the operator receive is reduced, complicating object pose estimation. A potential solution to this problem is the usage of 3D cameras. This would enhance the experience of the remote operator drastically by providing a sense of depth, texture, and even haptic feedback [6, 7].

III. VIRTUAL REALITY AND SYNTHETIC VIEWING

Virtual reality (VR) consists of an interactive computer simulation that enhances sensory, usually visual, feedback in such a way that the operator perceives the feeling of being immersed in the simulation [8]. VR applications in industry are various: simulation of mechatronic systems, prototype testing and training, assembly tasks assessment, to mention a few [9]. In the UKAEA, VR has been used as an approach trying to solve the lack of depth perception when performing remote operations and training inside JET. This consists in including a real-time VR replica of the teleoperated system interacting with JET's inner vessel (environment) on the operator's vision system [10] (Figure 3). Furthermore, the dynamic effects of VR simulation on the operators are currently being investigated and considered for improvement and verification of future reactors such as ITER [11]. Moreover, an extension of VR called Augmented Reality (AR), consisting in the superposition of virtual objects over a real scenario, is currently under development for applications on ITER [12].

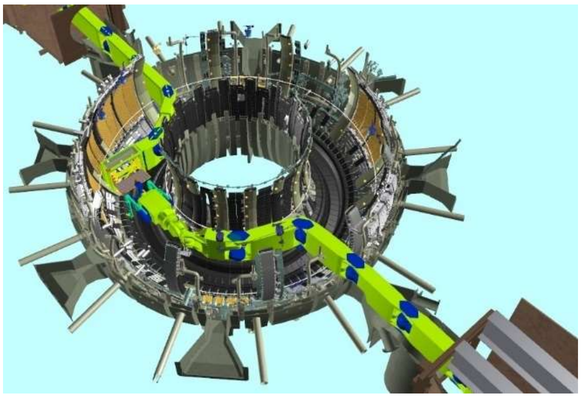


Figure 3. Simulation of two booms operating in the JET Tokamak.

The models of objects in VR and AR systems are required to encompass a description of its dynamic behaviour [8]. Such behaviour can be static or Newtonian. The first consist of stationary objects lacking interaction (visualisation purposes only) and the latter on objects that are ruled by real-world physics laws. In VR applications, these objects must be known in advance, which means that any sudden changes on the real scenario may not be captured in the virtual replica leading to potential collisions or unexpected behaviours. It is then necessary to utilise 3D feedbacked information to update or augment the virtual world accordingly.

Synthetic viewing is proposed as a tool to provide an enhanced environment state estimation. Synthetic viewing consists in combining data from the model and information gathered from multiple sensors to estimate the state of an environment with aims at real-time viewing. The fusion of this information allows to generate an up-to-date virtual environment, and due to the 3D nature of the information, it is possible to visualise the environment from multiple points of view avoiding occlusions and interferences typically present on video-only viewing systems. Although synthetic viewing is currently under assessment for applications in ITER [13], a large amount of research is being conducted to prove and demonstrate its capabilities for remote handling applications [14].

Recent developments on 3D cameras, also known as RGB-D sensors, allow to gather point clouds with millions of points with ~0.06mm precision at high rates. On one hand, the accuracy and density of this data can provide and outstanding insight of the remote environment, on the other hand, this data can be extremely large to achieve desirable transmission rates, particularly in remote operation.

Multiple approaches have been investigated to improve latency in point cloud data transmission. Some of them rely on performing object detection on the remote side (on-the-edge) aiming at transmitting only analytical parameters of the objects located on the scene. Methodologies vary from deep learning [15, 16] to classical fitting approaches such as RANSAC [17]. Computational time, detection accuracy and a priori model availability represent some of the drawbacks on this approach. For instance, in partially known scenarios such as decommissioning of gloveboxes, accurate 3D or analytical models of the objects are not always available.

Other techniques are focused on the compression of point cloud raw data using tree-like structures [14], video compression formats [18] among others [19]. The most important advantage of these methods is that there is no need for an a priori knowledge of the scene. However, even after compression, the transmitted data can still be large and bandwidth consuming. Moreover, compression may represent a loss in accuracy and resolution. To this end, a combination of object detection and point cloud data compression can represent a good trade-off between computation and transmission time, and quality.

The ultimate goal of the integration of synthetic viewing in teleoperation tasks, is to contribute in the development of a robust and reliable remote handling system by providing an intuitive vision system. This with aims at reducing training and operation times. In turn, this translates into shorter and more productive shutdown periods optimising operational costs.

IV. UK-ROK: SYNTHETIC VIEWS FOR A DIGITAL TWIN-BASED SYSTEM FOR REMOTE OPERATION

The UK-ROK project is a collaboration between the UK and the Republic of Korea (ROK) aiming to establish a digital-twin based teleoperation system operated between the two countries. This is a cross-institutions project having the University of Manchester and the Remote Applications in Challenging Environments (RACE) department of the UKAEA on the UK side, and the Korea Atomic Energy Research Institute (KAERI) on the ROK side.

Considering a traditional unilateral controlled teleoperation system, this project is focused on enhancing operator's performance by providing physical models of the remote environment, consisting of the remote manipulator and 3D cameras gathering information of their surroundings (Figure 4). The operator in front of the HID is set to use a digital twin that is replicating the remote manipulator and overlaying the point cloud data gathered from the remote environment on the real-time VR setup.

The information consisting, but not limited to, point cloud and robot joint positions, is then transmitted to the digital twin using the communications channel. Similarly, control commands (joint positions and velocities) are also send from the HID to the remote robot using the same communication channel. Hence, the communication is bidirectional.

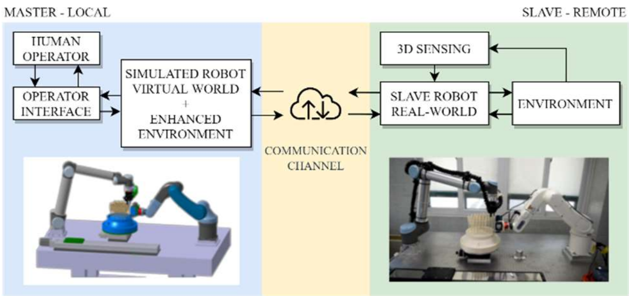
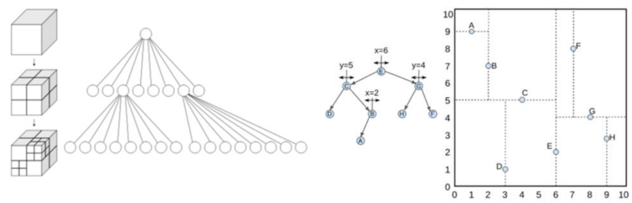


Figure 4. High-level teleoperation system architecture.

The communication channel is a virtual private network which tunnels the data transmitted from one end to the other over the internet. In terms of control, the framework of operation is expected to be: i) manual, ii) human planned and iii) human supervised. Due to limitations in this paper's scope, the control aspects of the manipulator will be kept to a minimum.

A. Point Cloud Data Compression using Tree-like Structures

As mentioned in the previous section, the efficient and fast transmission of point cloud data plays a key role in adequate functioning of the vision system. Hence, one of the most important research topics in this project relate to the compression, transmission, and decompression of point cloud data. In this research, the compression of point clouds is conducted using tree-based structure codification, more specifically octrees and kd-trees. Octrees and kd-trees perform a recursive decomposition which partitions space into cells of a determined size. Then, points lying within each cell are combined to generate a single data point. In the case of octrees, the size and number of cells is defined a priori, whereas on the kd-trees, the size is defined using hyperplanes defined by the median value computed for each axis recursively [14] (Figure 5).

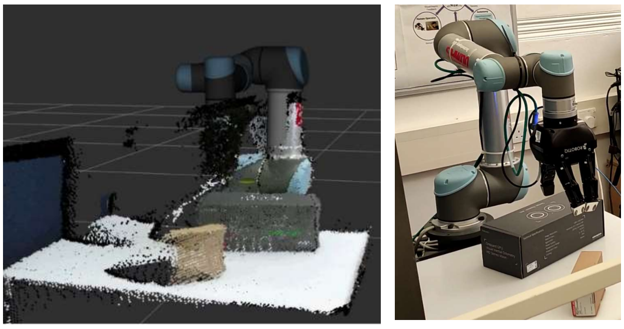


a) b)

Figure 5. Tree structures for point cloud data codification: a) octree and b) kd-tree [14].

B. 3D Enhanced Digital Twin for Improving Operators Performance

Initial point cloud data transmission experiments were conducted between two remote locations in the UK: Oxford and Manchester. The remote UR5 Robot manipulator with a RobotiQ Gripper for object manipulation (Figure 6b) is used for the teleoperation. The digital replica of the UR5 robot manipulator (Figure 6a) consisted in a VR representation. As can be observed in Figure 6a, with the integration of live point cloud data on the digital twin, the remote environment is visualised for the operator in a 3D fashion. Such enhancement allows the operator to perceive depth and colour, which improves the handling of objects. It is important to mention that prior to the integration of the 3D vision system onto the teleoperation system, the operator was only able to visualise image (video-like) feed from the remote environment. This created a burden to an adequate manipulation of objects and in some cases causing collisions and undesired behaviours.



a) b)

Figure 6. RGB-D enhanced master side. a) master side enhanced with live RGB-D data gathered from b) slave manipulator.

Initial tests showed that it is possible to update the reconstructed digital replica at a frame rate of ~11 frames per second, approximately one frame every 90 milliseconds. Even though this copes with expected update rates [13], the aim is to reduce even more this delay. To do so, we are currently investigating smarter ways to filter down the information we transmit to the remote side.

Object classification techniques based on deep learning appear to be a promising research path to identify relevant objects on the scene and to prioritise the transmission of those. KAERI has got a vast knowledge and experience in object segmentation from point cloud data and we are currently working on enhancing the proposed vision system. Another research path being explored relies on performing multiple levels of compression. This consists on dynamically variate the amount of data transmitted on a single frame based on the priorisation of elements on that frame. Likewise, we are focusing on a user-defined multiple-level compression system that would allow the user to manually prioritise the quality and transmission on certain objects on the scene. Both techniques aim at reducing the amount of data, hence reducing bandwidth consumption. With this, we predict and improvement on 3D data transmission times. Current work is expected to be integrated as a first trial on the decommissioning of contaminated gloveboxes in the near future.

V. Acknowledgements

This work is funded by the UK Engineering & Physical Sciences Research Council (EPSRC) Grant No. EP/S03286X/1. The authors would like to thank Michael Hellebrand and National Nuclear User Facility for Hot Robotics project (Grant No. EP/T011432/1) for leasing hardware to support this research.

VI. References

[1] F. Hu, "A Vision of an XR-Aided Teleoperation System toward 5G/B5G," in IEEE Communications Magazine, Vol 59, 2021.

[2] R. Buckingham, “Remote-handling challenges in fusion research and beyond”, Nature Phys 12, 391–393, 2016.

[3] B. Haist, “Remote handling preparations for JET EP2 shutdown”, Fusion Engineering and Design, 2009.

[4] UK Government, “Sellafield Contract Win for UK Firm”, Department of Climate Change and Energy, Online report, 2017.

[5] National Nuclear Laboratory, “Alpha Glovebox Decommissioning Feasibility Study”, UKRI Project 104068.

[6] F. Ryden, “Proxy method for fast haptic rendering from time varying point clouds”. In IEE/RSJ, San Francisco, USA, 2011.

[7] L. Niu, O. Suominen, “Eye-in-Hand Manipulation for Remote Handling: Experimental Setup”, in Proc. of MS&E, 2018.

[8] M. Matjaz, “Haptics for Virtual Reality and Teleoperation”, Series: Intelligent Systems, Control and Automation, Springer Dordrecht, 2012.

[9] D. Ma, X. Fan, J. Gausemeier and M. Grafe, “Virtual Reality & Augmented Reality in Industry”, Springer, Heidelberg, 2011.

[10] S. Sanders and A.C. Rolfe, “The use of virtual reality for preparation and implementation of JET remote handling operations”, Fusion Engineering and Design, Volume 69, 2003.

[11] C.J.M. Heemskerk, M.R. de Baar, et al, “Extending Virtual Reality simulation of ITER maintenance operations with dynamic effects”, Fusion Engineering and Design, 2011.

[12] ITER newsline, “Augmented reality - Assessing the future work environment”, https://www.iter.org/newsline/-/3509, 2020.

[13] ITER newsline 273, “The promises of ‘synthetic viewing’”, https://www.iter.org/newsline/273/1613, June 2013.

[14] S. Pacheco-Gutierrez, I. Caliskanelli and R. Skilton, “Point Cloud Compression and Transmission for Remote Handling Applications”, Journal of Software, Vol 16, Number 1, 2021.

[15] C. R. Qi, Li Yi, at al., “PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space”, In Proc. of the 31st Int. Conf. on Neural Information Processing Systems, USA, 2017.

[16] Y. Guo, H. Wang, Q. Hu, H. Liu, L. Liu and M. Bennamoun, "Deep Learning for 3D Point Clouds: A Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.

[17] M. A. Fischler and R. C. Bolles, “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography”. Comm. of the ACM, Vol 24, 1981.

[18] S. Schwarz, M. Preda, V. Baroncini, et al., “Emerging MPEG Standards for Point Cloud Compression”, IEEE Journal on Emerging and Selected Topics in Circuits and Systems 9, 2019.

[19] C. Cao, M. Preda, “3D Point Cloud Compression: A Survey”, 24th International Conference on 3D Web Technology, 2019.